

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 15:54:52 ; Search time 97.5789 Seconds
(without alignments)
221.222 Million cell updates/sec

Title: US-10-019-219-1

Perfect score: 912
Sequence: 1 TVVRLFLAMPQMWPCWLP.....WAACGARVKRRFLQTSLSR 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq_101002:*

1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	912	100.0	162	22	AB31701
2	912	100.0	166	22	AB31703
3	103.5	11.3	243	22	AAU46489
4	101.5	11.1	270	22	ABG30149
5	100	11.0	306	22	ABG19552
6	98	10.7	241	20	AAV73982
7	98	10.7	470	23	AAU91561
8	97	10.6	157	22	AAU49370
9	97	10.6	332	21	AAU83703
10	96.5	10.6	271	22	AAU58489

Result No.	Score	Query Match	Length	ID	Description
11	96.5	10.6	387	19	AAW81589
12	96	10.5	408	22	ABG59993
13	95.5	10.5	264	22	ABG05131
14	93.5	10.3	180	22	ABG12827
15	93.5	10.3	180	22	ABG14843
16	93.5	10.3	1061	20	AAW87504
17	93.5	10.3	1212	20	AAW87503
18	93	10.2	145	22	AAE06178
19	93	10.2	145	22	ABG34002
20	93	10.2	119	22	AAU58360
21	92.5	10.1	331	19	AAW81593
22	92.5	10.1	138	22	AAU64233
23	91.5	10.0	175	22	AAU87255
24	91	10.0	243	22	AAU60496
25	90.5	9.9	440	22	ABG05146
26	90.5	9.9	60	13	AAW24223
27	89.5	9.8	167	22	ABG19252
28	89.5	9.8	247	22	AAU43550
29	89	9.8	265	22	AAU43550
30	88.5	9.7	280	21	AAW34201
31	88.5	9.7	346	21	AAW04993
32	88.5	9.7	637	20	AAW26239
33	88.5	9.7	1232	23	AAW47961
34	88.5	9.7	106	23	AAU01922
35	88.5	9.6	207	22	AAU52687
36	88	9.6	233	22	AAW86095
37	88	9.6	1021	23	AAU79496
38	88	9.6	1170	23	AAU79500
39	88	9.6	1224	23	AAU15254
40	88	9.6	1224	23	AAU79497
41	88	9.6	1224	23	AAU80153
42	88	9.6	1224	23	AAU72891
43	88	9.6	1236	23	AAU98886
44	88	9.6			
45	88	9.6			

ALIGNMENTS

RESULT 1
ID AAB31701 standard; peptide: 162 AA.
XX AAB31701;
AC 30-APR-2001 (first entry)
XX
DT Peptide fragment of a human intestinal carboxylesterase (ICE).
XX
DE
XX
XX Intestinal carboxylesterase; ICE; tumour; cytotoxic T lymphocyte; CTL;
XX cytoxic factor; interleukin-2; interferon gamma; adenocarcinoma;
XX tumour necrosis factor; cancer; hepatocarcinoma; colon; kidney.
XX
XX Homo sapiens.
XX
XX WO200100784-A2.
XX
XX 04-JAN-2001.
XX
XX 27-JUN-2000; 2000WO-FR01791.
XX
XX 28-JUN-1999; 99FR-0008224.
XX
XX (INSTR) INSTR ROUSSY GUSTAVE.
XX
XX Ronsin C, Scott V, Triebel F;
XX WPI; 2001-112443/12.
XX
XX New peptides and its encoding nucleic acid derived from intestinal
XX carboxylesterase, useful as immunostimulants for treating cancer -

PS Claim 1; Page 3; 53p; French.
XX
CC The present sequence is derived from a human intestinal carboxylesterase
CC (ICE) polypeptide. ICE induces specific cytotoxic T lymphocytes
CC (CTL) and secretion by these CTL of cytotoxic factors, e.g.
CC Interleukin-2, interferon gamma and tumour necrosis factor. ICE
CC polypeptides and polynucleotides are used for treating cancer, by in
CC vivo or in vitro immunisation, particularly solid cancers and most
CC especially hepatocarcinoma or adenocarcinoma of colon and kidney. They
CC are also used to stimulate the immune system, and to increase, in
CC culture, the production of associated-associated CTL, for reinjection,
CC and/or to induce secretion of cytotoxic factors from CTL. Dendritic
CC cells loaded with ICE are used to induce such CTL in cultures.
XX
SQ Sequence 162 AA;
Query Match 100.0%; Score 912; DB 22; Length 162;
Best Local Similarity 100.0%; Pred. No. 5.3e-79;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TVVRLEFLAMPCMVPCMLPWRMTWSSSTAWVSWASSALETSQPATGATWTKMLHYA 60
DB 1 TVVRLEFLAMPCMVPCMLPWRMTWSSSTAWVSWASSALETSQPATGATWTKMLHYA 60
QY 61 GSSRISPTLEATLVSPFLASLRVARVCLRLCPYPKDSSTEPSRWVAMPSCPASLPAQ 120
DB 61 GSSRISPTLEATLVSPFLASLRVARVCLRLCPYPKDSSTEPSRWVAMPSCPASLPAQ 120
QY 121 LMSSPRMWPCTCLPYTKLTLRPWMAACGARVKRRFLQTLSTLSR 162
DB 121 LMSSPRMWPCTCLPYTKLTLRPWMAACGARVKRRFLQTLSTLSR 162
Db 121 LMSSPRMWPCTCLPYTKLTLRPWMAACGARVKRRFLQTLSTLSR 162
RESULT 2
AAB31703
ID AAB31703 standard; Protein: 166 AA.
XX
AC AAB31703;
XX
DT 30-APR-2001 (first entry)
XX
DE Protein encoded by an intestinal carboxylesterase (ICE) cDNA.
XX
KW Intestinal carboxylesterase; ICE; tumour; cytotoxic T lymphocyte; CTL;
KW cytotoxic factor; interleukin-2; interferon gamma; adenocarcinoma;
KW tumour necrosis factor; cancer; hepatocarcinoma; colon; kidney.
XX
OS Homo sapiens.
XX
PN WO200100784-A2.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-FR01791.
XX
PR 28-JUN-1999; 99FR-0008224.
XX
PA (INSR) INST ROUSSY GUSTAVE.
XX
PI Ronsin C, Scott V, Tiebel F;
XX
DR WPI: 2001-112443/12.
DR N-PSDB; AAF25258.
XX
PT New peptides and its encoding nucleic acid derived from intestinal
PT carboxylesterase, useful as immunostimulants for treating cancer -
XX
PS Disclosure; Fig 8a; 53p; French.
XX
CC The present sequence is encoded by the coding region of human intestinal
CC carboxylesterase (ICE) gene. ICE induces specific cytotoxic T
CC lymphocytes (CTL) and secretion by these CTL of cytotoxic factors,
CC e.g. Interleukin-2, Interferon gamma and tumour necrosis factor. ICE

CC polypeptides and polynucleotides are used for treating cancer, by in
CC vivo or in vitro immunisation, particularly solid cancers and most
CC especially hepatocarcinoma or adenocarcinoma of colon and kidney. They
CC are also used to stimulate the immune system, and to increase, in
CC culture, the production of associated-associated CTL, for reinjection,
CC and/or to induce secretion of cytotoxic factors from CTL. Dendritic
CC cells loaded with ICE are used to induce such CTL in cultures.
XX
SQ Sequence 166 AA;
Query Match 100.0%; Score 912; DB 22; Length 166;
Best Local Similarity 100.0%; Pred. No. 5.5e-79;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TVVRLEFLAMPCMVPCMLPWRMTWSSSTAWVSWASSALETSQPATGATWTKMLHYA 60
DB 5 TVVRLEFLAMPCMVPCMLPWRMTWSSSTAWVSWASSALETSQPATGATWTKMLHYA 64
QY 61 GSSRISPTLEATLVSPFLASLRVARVCLRLCPYPKDSSTEPSRWVAMPSCPASLPAQ 120
DB 65 GSSRISPTLEATLVSPFLASLRVARVCLRLCPYPKDSSTEPSRWVAMPSCPASLPAQ 124
QY 121 LMSSPRMWPCTCLPYTKLTLRPWMAACGARVKRRFLQTLSTLSR 162
DB 125 LMSSPRMWPCTCLPYTKLTLRPWMAACGARVKRRFLQTLSTLSR 166
Db 125 LMSSPRMWPCTCLPYTKLTLRPWMAACGARVKRRFLQTLSTLSR 166
RESULT 3
AAU46489
ID AAU46489 standard; Protein: 243 AA.
XX
AC AAU46489;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #7385.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteophtic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
XX
PR 02-JUN-2000; 2000US-208841P.
XX
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIYA CORP.
XX
PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhattacha A;
XX
DR L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI: 2001-616774/71.
XX
DR N-PSDB; AAS59532.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 7684; 1069p; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis), acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.

P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by enzyme linked immunosorbent assay (ELISA).

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 243 AA:

Query Match 11.3%; Score 103.5; DB 22; Length 243;
Best Local Similarity 28.1%; Pred. No. 0.055;
Matches 47; Conservative 13; Mismatches 62; Indels 45; Gaps 9;

OY 8 AW-LPGMWY---PCWLP-NRTWMSSTAWVMSASALETS-TOPATGATWTKMLHYAG 61
DB 6 AMSSTKMSYTRPGMWEVTRTPMRSSRRAMVARSAYFDPAATPLAGV----- 55
OY 62 SSRISPLEATLVSPFLA-----SLRVAVCLRLCPYPKDSSTEPSWRVAMPSCPAS 116
DB 56 --RRNDPELIMMTSPMLARSSGSTRVRSAATAMTLISNTRHSSR--GNEPISPGDVAL 111
OY 117 LPAQLMSSPRW-WPTCLPY-----TKLTRPPW 143
DB 112 TPALLIRASHMSWETAIVAVASSVSTKGEALGSDTLTSTASW 158

RESULT 4

ABG30149
ID ABG30149 standard; Protein; 270 AA.

AC ABG30149;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #30140.

KM Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PADB; AAS94336.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 20; SEQ ID No 60508; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome CC and gene mapping, and in recombinant production of (II). The CC polynucleotides are also used in diagnostics as expressed sequence tags CC for identifying expressed genes. (II) is useful in gene therapy techniques CC to restore normal activity of (II) or to treat disease states involving CC (II). (II) is useful for generating antibodies against it, detecting or CC quantifying a polypeptide in tissue, as molecular weight markers and as CC a food supplement. (II) and its binding partners are useful in medical CC imaging of sites expressing (II). (I) and (II) are useful for treating CC disorders involving aberrant protein expression or biological activity. CC The polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. ABG00010-ABG30377 represent novel human CC diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 270 AA:

Query Match 11.1%; Score 101.5; DB 22; Length 270;
Best Local Similarity 23.3%; Pred. No. 0.097;
Matches 34; Conservative 10; Mismatches 45; Indels 57; Gaps 9;

OY 9 WLCMWYPCGLPRTW-----WSSSTAWVMSASALETSQPTGATWTKMLHYAG 61
DB 158 WRSCTQRMWLCWPHWTLCTQSWW---LCWGHWRSC-----WTQSWMLCWPW-- 203
OY 62 SSRISPLEATLVSPFLASLRVAVCLR---LLCPYPKDSSTEPSWRVAMPSCPASLP 118
DB 204 -----RLCWTSQSWMLCWPWHTMNSCTQ-SWMLCWPW----- 233

OY 119 AOLMSSPRWMPCLPYTKL-TLRPW 143

DB 234 -RLCWSSQSWW-LCPHWRLYWTSWW 257

RESULT 5

ABG19552
ID ABG19552 standard; Protein; 306 AA.

AC ABG19552;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #19543.

KM Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PADB; AAS83739.

PT New isolated polynucleotide and encoded polypeptides, useful in

diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

Claim 20; SEQ ID No 49911; 103bp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 306 AA;

Query Match 11.0%; Score 100; DB 22; Length 306;
Best Local Similarity 26.0%; Pred. No. 0.16; Mismatches 40; Conservative 15; Mismatches 57; Indels 42; Gaps 9;

Db 10 LFCMAYPCMLP-----W-RTWM-----SSSTAWVSSASLETSTPATGATW 53
11 LFGQKSHMWSKIRPAMWTRSMALPQAPLSGASWPSFVRT--TPTTGTCT 168

QY 54 TKMLHYAGSSRSPLETLTVSPFLASIR-----VARVCLRLCPRYPKDSSTE 103
Db 169 PSWMSWRSTSRSSSL-----NPTESLKNLSRTSKPFGWSQAMAPCPILPPASRYE 222

QY 104 PSWRYAWPSCPASLPQAQLMSSPRMPTCLPVTKL 137
Db 223 -LWPEWTW-----LGGHATSSQW--KCVPERKL 248

RESULT 6
AAV73982
ID AAV73982 standard; Protein: 241 AA.

AC AAV73982;
DT 14-MAR-2000 (first entry)

DE Human prostate tumor EST fragment derived protein #169.

KW Pancreas; tumor; EST; expressed sequence tag; human; cytosolic;
KW treatment.

OS Homo sapiens.
PN DE19820190-A1.
PD 04-NOV-1999.
PE 28-APR-1998; 98DE-1020190.
PR 28-APR-1998; 98DE-1020190.
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;

WP1; 1999-621386/54.
DR N-PSDB; AA252913.

New human nucleic acid sequences from pancreatic tumors, and related proteins.

Claim 23; Page 380; 502pp; German.

This invention describes novel polypeptides and their encoding nucleic acids derived from human pancreatic tumor tissue which have cytostatic activity. The sequences are also useful in producing pharmaceutical compositions for treatment of pancreatic tumors. AAV73814-V74252 represent protein fragments encoded by the human pancreatic tumor cDNA library derived expressed sequence tag (EST) sequences represented in AA252858-253014.

Sequence 241 AA;

Query Match 10.7%; Score 98; DB 20; Length 241;
Best Local Similarity 25.7%; Pred. No. 0.18; Mismatches 45; Conservative 10; Mismatches 64; Indels 56; Gaps 9;

QY 9 WPCMM-----VPCMLPRTWMMSSSTAWVSSASLETSTPATGATGATW 58
Db 38 WPAALRNKVEAPERMSPPWCPW-AMCWQ-----WEPW-----LWGMEPGTGTSTEFQS 87

QY 59 YA-----GSSRISPTLEATLTVSPFLASLVARVCLRLCPY-----PKDSSTE 103
Db 88 EATGQTACQTSRTGNGLEPMTWELRLRRHPSSEKSLPLRAPQPKNPRKQKH 147

QY 104 PSWRYAWP-----SCPASLPQAQLMSSPRMPTCLPVTKLRLPWWAAGARVKKRFL 155
Db 148 PRRKPRMPKTKSC-----SSPAPMP-----RPRTAPRKPRCRRL 184

RESULT 7
AAU91561
ID AAU91561 standard; Protein: 470 AA.

AC AAU91561;
DT 02-JUL-2002 (first entry)

DE Amino acid sequence of human PHOR1-F5D6 splice variant C (frame 2).

KW Human; PHOR1-F5D6; prostate cancer; cytosolic.

OS Homo sapiens.
PN WO200214501-A2.
PD 21-FEB-2002.
PE 17-AUG-2001; 2001WO-US25862.
PR 17-AUG-2000; 2000US-226241P.
PA (AGEN-) AGENSYS INC.
PI Hubert RS, Raitano AB, Faris M, Challita-Eld PM, Ge W;
PI Jakobovits A;
DR WP1; 2002-269193/31.

Monitoring PHOR1-A1/PHOR1-F5D6 gene products for monitoring presence of cancer in subject, by determining status of PHOR1-A1/PHOR1-F5D6 gene products in tissue sample from subject and comparing it to normal sample.

Example 45; Page 217; 250pp; English.

The present invention relates to the isolation of novel human genes

PT for modulating thermogenesis in tissues, e.g. for treating obesity
PT or muscle wasting caused by infection or cancer
PS Disclosure: Fig 1A-C; 98pp; English.
XX
XX
CC Sequences AAW81588 to AAW81590 represent protein fragments encoded by
CC the three reading frames of the human uncoupling protein 3 (UCP3) gene.
CC The invention provides human and mouse UCP3 genes (AAV7110 and
CC AAV7112) encoding UCP3 proteins (AAW81587 and AAW81595) respectively. A
CC host cell transformed with a construct comprising the UCP3 nucleic acid
CC can be used for the recombinant production of the protein. The UCP3 is
CC involved in the regulation of thermogenesis in mammals. The nucleic acids
CC (AAV7110 to AAV7112) can be used for identifying compounds which alter
CC UCP3 activity. Enhancers of UCP3 can be used for enhancing protein
CC catabolism in a mammal and can be used as anti-obesity drugs. Inhibitors
CC of UCP3 can be used for inhibiting protein catabolism in a mammal such as
CC inhibiting muscle wasting. They can be used for curtailing muscle wasting
CC due to infection (e.g. HIV, cancer, tumour cachexia, muscle diseases
CC (e.g. muscular dystrophy) or as a possible treatment for non-insulin
CC dependent diabetes mellitus. The products can also be used for detection
CC and diagnosis.
XX
XX
SQ Sequence 387 AA:

Query Match 10.6%; Score 96.5; DB 19; Length 387;
Best Local Similarity 22.7%; Pred. No. 0.44;
Matches 47; Conservative 17; Mismatches 60; Indels 83; Gaps 11;

OY 12 GAAVGCWL--PW-----RTWMMSSSTAWMSALETSS-----QDPAT- 49
DB 114 CSTVACGMAPSPCGLRVPAAPTMGMPACASARSPSASMTSPSRCTPPKARTTPASL 173
OY 50 -----GATWTKN-----LHYAGSSRSISPTLEATLVSPFLASLRVR 86
DB 174 PGFWPAAPOEPRPVPSPQWMSRDRPATYSGHPATENTAGTMTTTEPSP---GKKEG 230
OY 87 VCLRLICPPYPRDSSTPSPW-----RVAM-----PSCPA--SLPAQIMSSPRMPTCL 132
DB 231 AGCKELCPTSGMISSTVLRLMPTTSSRRSCMTTCSILTSPATLSLPLEPASVQWMP--- 287
OY 133 PYTKLTLRP-----WV 143
DB 288 --PRMTWRPGITHLQASTASAPSTVRWV 312

RESULT 12
ABG59993
ID ABG59993 standard; Protein; 408 AA.
XX
XX
AC ABG59993;
XX
XX
DT 30-JUL-2002 (first entry)
XX
XX
DE Human DITHP polypeptide #51.
XX
XX
XX Human; DITHP; diagnostic and therapeutic polypeptide; bone; testis; skin;
XX cell proliferative disorder; cancer; tumour; autoimmune disorder; brain;
XX inflammatory disorder; viral infection; bacterial infection; seizure;
XX fungal infection; parasitic infections; developmental disorder; breast;
XX endocrine disorder; metabolic disorder; neurological disorder; cervix;
XX gastrointestinal disorder; transport disorder; gene therapy; kidney;
XX adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;
XX thymus.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200220754-A2.
XX
XX
PD 14-MAR-2002.
XX
XX
PF 29-AUG-2001; 2001WO-US27127.
XX
XX
PR 05-SEP-2000; 2000US-229747P.

PR 05-SEP-2000; 2000US-229748P.
PR 05-SEP-2000; 2000US-229749P.
PR 05-SEP-2000; 2000US-229750P.
PR 05-SEP-2000; 2000US-229751P.
PR 05-SEP-2000; 2000US-230583P.
PR 06-SEP-2000; 2000US-230585P.
PR 06-SEP-2000; 2000US-230514P.
PR 06-SEP-2000; 2000US-230515P.
PR 06-SEP-2000; 2000US-230517P.
PR 06-SEP-2000; 2000US-230518P.
PR 06-SEP-2000; 2000US-230519P.
PR 06-SEP-2000; 2000US-230595P.
PR 06-SEP-2000; 2000US-230597P.
PR 06-SEP-2000; 2000US-230598P.
PR 06-SEP-2000; 2000US-230599P.
PR 06-SEP-2000; 2000US-230610P.
PR 06-SEP-2000; 2000US-230656P.
PR 06-SEP-2000; 2000US-230988P.
PR 07-SEP-2000; 2000US-230951P.
PR 07-SEP-2000; 2000US-231163P.
PR 07-SEP-2000; 2000US-231167P.
XX
XX
PA (INCYTE) INCYTE GENOMICS INC.
XX
XX
PI Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;
PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;
PI Monyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
PI Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;
PI Marwaha R, Chen AJ, Chang SC, Au AP, Imman RR;
XX
XX
DR WPI: 2002-383054/41.
DR N-PSDB: ABR71585.
XX
XX
PT An isolated polynucleotide useful in diagnostics and therapeutics -
XX
XX
PS Claim 29; Page 561-562; 686pp; English.
XX
XX
XX The invention relates to human diagnostic and therapeutic (dithp)
XX polynucleotides and their associated polypeptides (DITHP polypeptides).
XX The sequences of the invention are used in the treatment and diagnosis of
XX cell proliferative disorders (e.g. atherosclerosis, cirrhosis), cancers
XX (e.g. tumours of the adrenal gland, bone, bone marrow, breast,
XX cervix, kidney, lung, ovary, pancreas, prostate, skin, spleen, testis or
XX thymus), autoimmune/inflammatory disorders (e.g. asthma, bronchitis,
XX psoriasis, osteoporosis), viral infections, bacterial infections, fungal
XX infections, parasitic infections, developmental disorders (e.g. anaemia,
XX epilepsy), seizure disorders (e.g. cerebral palsy, spina bifida),
XX endocrine disorders (e.g. thrombosis, aneurysm), metabolic disorders
XX (e.g. obesity, diabetes), neurological disorders (e.g. stroke,
XX amyotrophic lateral sclerosis, multiple sclerosis), gastrointestinal
XX disorders (e.g. ulcerative colitis, lysinuria) and transport disorders
XX (e.g. myotonic dystrophy, catatonias, peripheral neuropathy). Sequences
XX ABG59943-ABG60220 represent human DITHP polypeptides of the invention.
XX
XX
SQ Sequence 408 AA:

Query Match 10.5%; Score 96; DB 23; Length 408;
Best Local Similarity 25.0%; Pred. No. 0.52;
Matches 42; Conservative 16; Mismatches 60; Indels 50; Gaps 10;

OY 8 AMLPCMVPCMLPWT-----WMMSSST-----AWMSALETSTROPAT 49
DB 5 SMASVATAPCW--MATRASPTVPCWPRASSTISSGNGPRAVAURCGSTATSRPSP 62
OY 50 GATWTKMLHAYAGSSRSISPTLEATLVSPFLASLRVAVRCLRLCPTPKDSTSPSRVA 109
DB 63 AAYWTSCTRAAMTWRSILP--VEDVLAASYLHMVDIVKVGKGR---QEKRSIDPG---- 114
OY 110 WPCSPASLPQAQIMSSPRMPTCLPVTKTLRLRPV--NAA--CGARVRR 153
DB 115 -NPAPGAERPAQ-----PFC-----PWPVWTLADLCPARAKR 144

RESULT 13
ID ABG05131 standard; Protein: 264 AA.
XX
AC ABG05131;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #5122.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
OS Homo sapiens.
XX
PN WO200175067-A2.
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS69318.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID NO 35490; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 264 AA;
XX
Query Match 10.5%; Score 95.5; DB 22; Length 264;
Best Local Similarity 31.2%; Pred. No. 0.35;
Matches 29; Conservative 10; Mismatches 41; Indels 13; Gaps 3;

RESULT 14
ID ABG12827 standard; Protein: 180 AA.
XX
AC ABG12827;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #12818.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
OS Homo sapiens.
XX
PN WO200175067-A2.
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS77014.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID NO 43186; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 180 AA;
XX
Query Match 10.3%; Score 93.5; DB 22; Length 180;
Best Local Similarity 27.1%; Pred. No. 0.35;
Matches 39; Conservative 8; Mismatches 48; Indels 49; Gaps 8;

Db	92	WC--AGTGPPAATSPRAFTPTTSPSCSPASTASP-----TFCSTASSARPTGTWPAS	143
QY	109	---AMPSCPASLPAQOLMSSPRMWP	129
Db	144	AGPAMPSSPAGPAGPGRPTRMWP	167
RESULT 15			
ABG14843	ID	ABG14843 standard; Protein; 180 AA.	
AC	ABG14843;		
XX	18-FEB-2002	(first entry)	
DE	Novel human diagnostic protein #14834.		
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;		
KW	food supplement; medical imaging; diagnostic; genetic disorder.		
XX			
OS	Homo sapiens.		
XX	WO200175067-A2.		
PN	11-OCT-2001.		
PD	30-MAR-2001; 2001WO-US08631.		
XX			
PE	31-MAR-2000; 2000US-0540217.		
PR	23-AUG-2000; 2000US-0649167.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Drmanac RT, Liu C, Tang YT;		
XX	WPI: 2001-639362/73.		
DR	N-PSDB; NMS79030.		
XX			
PT	New isolated polynucleotide and encoded polypeptides, useful in		
PT	diagnostics, forensics, gene mapping, identification of mutations		
PT	responsible for genetic disorders or other traits and to assess		
PT	biodiversity -		
PS	Claim 20; SEQ ID NO 45202; 103bp; English.		
XX			
CC	The invention relates to isolated polynucleotide (I) and		
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,		
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome		
CC	and gene mapping, and in recombinant production of (II). The		
CC	polynucleotides are also used in diagnostics as expressed sequence tags		
CC	for identifying expressed genes. (I) is useful in gene therapy techniques		
CC	to restore normal activity of (II) or to treat disease states involving		
CC	(II). (II) is useful for generating antibodies against it, detecting or		
CC	quantitating a polypeptide in tissue, as molecular weight markers and as		
CC	a food supplement. (I) and its binding partners are useful in medical		
CC	imaging of sites expressing (II). (I) and (II) are useful for treating		
CC	disorders involving aberrant protein expression or biological activity.		
CC	The polypeptide and polynucleotide sequences have applications in		
CC	diagnostics, forensics, gene mapping, identification of mutations		
CC	responsible for genetic disorders or other traits to assess biodiversity		
CC	and to produce other types of data and products dependent on DNA and		
CC	amino acid sequences. ABG00010-ABG30377 represent novel human		
CC	diagnostic amino acid sequences of the invention.		
CC	Note: The sequence data for this patent did not appear in the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pct_sequences.		
XX			
XX	Sequence 180 AA;		

```

Oy      25  MWSSSTAWVWMASSALETSQNPATGAT-----WTK-----55
Db      34  MWASSSPACCCRPPL--RASCAPAGANAIPRAARTSSGCGASPMWSSSPASCPPTCCC 91
Oy      56  WLHVAGSSR---ISPTEALITVSPFLASLVARVVCILCPYKDDSTEP---SWRV-108
Db      92  WC--AGTRPAATSPRASTITTTSPSCSPASTASP-----TPCSTASSARPPIGTWPAS 143
Oy      109 ---AMPSCPASLPAOLMSSPRWMP 129
Db      144 AGPAWPSSPAGCPAGGKPTWMP 167

```

Search completed: March 27, 2003, 16:10:32
Job time : 99.5789 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 16:08:52 ; Search time 35.0526 Seconds
(without alignments)
135.982 Million cell updates/sec

Title: US-10-019-219-1
Perfect score: 912
Sequence: 1 TVVRLFLAMPCMVPCWLP.....WAGCARVKRRFLQTSLSR 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTRUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	9.6	1224	4	US-09-930-872-4
2	87.5	9.6	1422	4	US-08-469-260A-81
3	87	9.5	449	2	US-08-819-458A-16
4	84.5	9.3	498	2	US-08-660-963-12
5	80	8.8	539	2	US-08-735-041A-2
6	80	8.8	539	3	US-09-190-476B-2
7	80	8.8	539	3	US-09-190-889A-2
8	80	8.8	539	3	US-09-190-938B-2
9	80	8.8	539	5	PCF-US95-09261-2
10	79.5	8.7	119	2	US-08-652-816A-10
11	79.5	8.7	142	2	US-08-480-774A-2
12	79	8.7	117	3	US-08-545-809A-114
13	79	8.7	278	4	US-09-260-527-3
14	76.5	8.4	76	4	US-08-851-362D-22
15	76.5	8.4	118	3	US-08-545-809A-142
16	76.5	8.4	516	2	US-08-794-216-3
17	76.5	8.4	516	2	US-08-749-289-3
18	76.5	8.4	1711	2	US-08-342-930-2
19	76	8.3	187	2	US-08-078-311-2
20	76	8.3	187	2	US-08-460-402-2
21	76	8.3	209	4	US-09-109-100-8
22	76	8.3	209	4	US-09-109-100-9
23	76	8.3	209	4	US-09-109-100-11
24	76	8.3	209	4	US-09-109-100-12
25	76	8.3	209	4	US-09-109-100-13
26	76	8.3	209	4	US-09-109-100-14
27	76	8.3	209	4	US-09-109-100-15

28	76	8.3	209	4	US-09-109-100-17	Sequence 17, Appl
29	76	8.3	209	4	US-09-109-100-18	Sequence 18, Appl
30	76	8.3	212	4	US-09-109-100-10	Sequence 10, Appl
31	76	8.3	235	1	US-08-243-545-6	Sequence 6, Appl1
32	76	8.3	235	2	US-08-993-962-6	Sequence 6, Appl1
33	76	8.3	235	4	US-09-160-841-6	Sequence 6, Appl1
34	76	8.3	235	4	US-09-109-100-1	Sequence 1, Appl1
35	76	8.3	235	5	PCF-US94-05365-6	Sequence 6, Appl1
36	76	8.3	276	4	US-09-094-964-3	Sequence 3, Appl1
37	76	8.3	635	1	US-08-184-327A-4	Sequence 4, Appl1
38	76	8.3	635	2	US-08-078-311-1	Sequence 1, Appl1
39	76	8.3	635	2	US-08-460-402-1	Sequence 1, Appl1
40	76	8.3	635	5	PCF-US95-00670-4	Sequence 4, Appl1
41	75.5	8.3	206	2	US-08-438-439C-19	Sequence 19, Appl
42	74.5	8.2	98	1	US-08-478-039-75	Sequence 75, Appl
43	74.5	8.2	116	3	US-08-476-349A-75	Sequence 75, Appl
44	74.5	8.2	116	3	US-08-545-809A-140	Sequence 140, Appl
45	74.5	8.2	119	4	US-09-025-769B-39	Sequence 39, Appl

ALIGNMENTS

```
RESULT 1
US-09-930-872-4
; Sequence 4, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Fridlie, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. 6448388e1 Human Proteases and Polynucleotides Encoding the
; FILE REFERENCE: Lex-0219-USA
; CURRENT APPLICATION NUMBER: US/09/930,872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1224
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-930-872-4

Query Match          9.6%; Score 88; DB 4; Length 1224;
Best Local Similarity 25.5%; Pred. No. 1.3;
Matches 37; Conservative 18; Mismatches 56; Indels 34; Gaps 7;

QY 12 CMNVPCWLPWRTWMSSTPRAWVSWASALETSTOPATGATWTKWLYAGSSRISPTLEA 71
Db 1042 CLDGRCHKPKKQWLVSV--AW-SQCSVTCERGTOKRFLLCAEKYV--SGKYRELAASKC 1095

QY 72 TLVSPPLASLRARVCRLICPPYKDSSTERSMRYAMSCPASLPDQIMSSBRMWPPTC 131
Db 1096 SHLPKP---SLIERACAPLPCPRHPPFAAGPS-RGSMFASP-----WSOC 1138

QY 132 LPVTKLIRPMWMAAGARVKRRFLQ 156
Db 1139 T-----ASCGGGVQTRSYQ 1152

RESULT 2
US-08-469-260A-81
; Sequence 81, Application US/08469260A
; Patent No. 6451578
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMM J. PILOT-MATTIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MOERHOFF
```

```

APPLICANT: JAMES C. ERKER
APPLICANT: SHERI L. BUIJK
APPLICANT: ISA K. MUSHANAR
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
NUMBER OF SEQUENCES: 716
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,260A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/424,550
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FOREMBSTI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5527 .PC.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 81:
SEQUENCE CHARACTERISTICS:
LENGTH: 1422 amino acids
TYPE: amino acid
STRANDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-260A-81

Query Match          9.6%; Score 87.5; DB 4; Length 1422;
Best Local Similarity 21.1%; Pred. No.1.7;
Matches 45; Conservative 18; Mismatches 55; Indels 95; Gaps 9.

QY      10 LPCHVWPCMLPWRTVMWSSSSTAWSMASALETSTOPATGATWTKWLHYACSSRI-SPT 68
       ||::||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      594 LPVLSPQMILKRKKWRBSVHNS-----PPMRPWLQLTSXKVQSFG 634

QY      69 LEAT-----LTVSDFL-----ASLRVAR-----VCLRLICEPY--- 96
       | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      635 LVLSHWKPPLKNLPFLGLMQSLLSXSIAVAKSLYLTIPLHHACLILSRVLPHTLR 694

QY      97 -----PRDSSTPSRWAMP-----SCRAS----- 116
       | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      695 SKCSCHYLEAOLRPSLOLTLEXHKRSXMGLEKOLLVHGHRWVLSLTCAAMLAHPDLAX 754

QY      117 -LPAQLMSSPRMWPFCLPVTKLTPR--WNAAC 146
       | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      755 HLNAAKWVGXLMISLVASTPRSIRPQELMASC 787

RESULT 3
US-08-819-458A-16
Sequence 16, Application US/08819458A
Patent No. 5891669
GENERAL INFORMATION:
APPLICANT: Jensen, Ejner B.
APPLICANT: Cherry, Joel
APPLICANT: Elrod, Susan L.
TITLE OF INVENTION: Methods for Producing Polypeptides
TITLE OF INVENTION: In Respiratory-Deficient Cells
NUMBER OF SEQUENCES: 20
```

```

CORRESPONDENCE ADDRESS:
ADDRESSSEE: No. 58916690 No. 5891669disk of No. 5891665th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: U.S.A.
ZIP: 10174

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/819,458A
FILING DATE: 17-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 5215,000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:

INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-819-458A-16

Query Match          9.5%; Score 87; DB 2; Length 449;
Best Local Similarity 20.6%; Pred.No. 0.46;
Matches 41; Conservative 18; Mismatches 60; Indels 80; Gaps 7,
QY 22 RTWW-----WSSSRTAWYSW-----ASSALETSTPATGAT 52
      |   ||| :   |   |   |   |   |   |   |   |   |   |
Db 114 RKWFSSIMVSWTLRSPOLYLEFMSPRLLHSNOFIACADILLPOLRRSVILDTSTVPL--LS 171
      |   |   |   |   |   |   |   |   |   |   |   |   |
QY 53 WTKWLHVAGSSRIISPTLEATLIVSPFLASLRVARVCLR-----90
      |   |   |   |   |   |   |   |   |   |   |   |   |
Db 172 WKSLTMECTDLTEOVWOSTLTWTISMLPKIRSTRAVLREPMTESILSPYLMARPDIVSGA 231
QY 91 -LLCP-----PYPKDST-----EPSWRV-----AMPSCPASLPALQMS 123
      |   |   |   |   |   |   |   |   |   |   |   |   |
Db 232 TLLODLRLHLTSAPSPPLASSSPRCPPRWLVQATLLSSTRLVARAATASCSCCTPARSQ 291
QY 124 SPRWPFTCLPYTKTLRPW 142
      |   |   |   |   |   |   |   |   |   |   |   |   |
Db 292 LRSWIIFLEPPIPSFRSM 310

RESULT 4
US-08-660-963-12
Sequence 12, Application US/08660963
Patent No. 5852187
GENERAL INFORMATION:
APPLICANT: Thorne, Michael O.
APPLICANT: Gaylinn, Bruce D.
APPLICANT: Horikawa, Reiko
APPLICANT: Lyons Jr., Charles E.
TITLE OF INVENTION: MOLECULAR CLONING OF THE OVINE PITUITARY
TITLE OF INVENTION: GROWTH HORMONE RELEASING HORMONE RECEPTOR
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESSES:
ADDRESSEE: POPHAM, HAIR, SCHNOBRICH & KAUFMAN, LTD.
STREET: Metropolitan Square Building, Suite 800, 1450
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/660,963
; FILING DATE: 12-JUN-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Shaughnessy, Brian P.
; REGISTRATION NUMBER: 32,747
; REFERENCE/DOCKET NUMBER: 18046.036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-824-8000
; TELEFAX: 202-824-8199
; TELEX: 248516
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 498 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-660-963-12

```

```

Query Match          9.3%; Score 84.5; DB 2; Length 498;
Best Local Similarity 26.6%; Pred. No. 0.93;
Matches 41; Conservative 14; Mismatches 40; Indels 59; Gaps 9;

```

```

QY 20 PRTW-----WSSSTAN-VSMASALSTQPA-----TGATWT 54
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 179 PMATSRLOPWPSS--WSLSGSTAPGTSTPSCSPPLSRQLCSRTPPSFTGRTWT 236
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 55 KMLHYAGSRISPTLEATLTVPFLASLRVARYCLTLPYPKDSTEPSW-RVAMPSC 113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 237 -----TAASPLS-----CARLLPPLISRPPTSAGCWMKICTPAS 270
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 114 PASLPAQLMSSPRMPTCLPVTKLIRPWWACG 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 271 PPHCPAQGGSSGGWF--SLPGGFLCSSP---ACG 299
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 5
; US-08-735-041A-2
; Sequence 2, Application US/08735041A
; Patent No. 5914251
; GENERAL INFORMATION:
; APPLICANT: Farrell, Catherine L.
; APPLICANT: Martin, Francis H.
; APPLICANT: Yabkowitz, Rachel
; TITLE OF INVENTION: PLACENTAL-DERIVED PROSTRATE GROWTH
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/735,041A
; FILING DATE: 22-OCT-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657

```

```

; REFERENCE/DOCKET NUMBER: A-414
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 805.447.4112
; TELEFAX: 805.499.6751
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 539 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-735-041A-2

```

```

Query Match          8.8%; Score 80; DB 2; Length 539;
Best Local Similarity 28.8%; Pred. No. 3;
Matches 17; Conservative 9; Mismatches 11; Indels 22; Gaps 3;

```

```

QY 17 CWLP-----WRTMWSSSTAWSWASSALSTQPATGATWT---KMLHYAG 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 398 CWLSPGLRRTASGALWRSGLMAAAPSSW-----STSLRPAQGGPSWGRGCOMQLQVLG 448
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 6
; US-09-190-476B-2
; Sequence 2, Application US/09190476B
; Patent No. 6025204
; GENERAL INFORMATION:
; APPLICANT: Farrell, Catherine L.
; APPLICANT: Martin, Francis H.
; APPLICANT: Yabkowitz, Rachel
; TITLE OF INVENTION: PLACENTAL-DERIVED PROSTRATE GROWTH
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/190,476B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/735,041
; FILING DATE: 22-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-414
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 805.447.4112
; TELEFAX: 805.499.6751
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 539 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-190-476B-2

```

```

Query Match          8.8%; Score 80; DB 3; Length 539;
Best Local Similarity 28.8%; Pred. No. 3;
Matches 17; Conservative 9; Mismatches 11; Indels 22; Gaps 3;

```

```

QY 17 CWLP-----WRTMWSSSTAWSWASSALSTQPATGATWT---KMLHYAG 61

```

Db 398 CWLSPGLRTASGALWMSGAWAAPSSW-----STSLRPGGSPWTGCGWLQVIG 448

RESULT 7

US-09-190-889A-2
Sequence 2, Application US/09190889A
Patent No. 6075008
GENERAL INFORMATION:
APPLICANT: Farrell, Catherine L.
APPLICANT: Martin, Francis H.
APPLICANT: Yabkowitz, Rachel
TITLE OF INVENTION: PLACENTAL-DERIVED PROSTATE GROWTH
TITLE OF INVENTION: FACTOR
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 De Havilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/190,889A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/735,041
FILING DATE: 22-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Mazza, Richard J.
REGISTRATION NUMBER: 27,657
REFERENCE/DOCKET NUMBER: A-414
TELECOMMUNICATION INFORMATION:
TELEPHONE: 805.499.6751
TELEFAX: 805.499.6751
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-190-889A-2

Query Match 8.88; Score 80; DB 3; Length 539;
Best Local Similarity 28.88; Pred. No. 3;
Matches 17; Conservative 9; Mismatches 11; Indels 22; Gaps 3;
OY 17 CWLP-----WRTWWSSTAVSWASSALESTQPGATWT---KMLHYAG 61
Db 398 CWLSPGLRTASGALWMSGAWAAPSSW-----STSLRPGGSPWTGCGWLQVIG 448

RESULT 8

US-09-190-938B-2
Sequence 2, Application US/09190938B
Patent No. 6197939
GENERAL INFORMATION:
APPLICANT: Farrell, Catherine L.
APPLICANT: Martin, Francis H.
APPLICANT: Yabkowitz, Rachel
TITLE OF INVENTION: PLACENTAL-DERIVED PROSTATE GROWTH
TITLE OF INVENTION: FACTOR
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 De Havilland Drive

CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/190,938B
FILING DATE: 12-No. 6197939-1998
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mazza, Richard J.
REGISTRATION NUMBER: 27,657
REFERENCE/DOCKET NUMBER: A-414
TELECOMMUNICATION INFORMATION:
TELEPHONE: 805.499.6751
TELEFAX: 805.499.6751
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-190-938B-2

Query Match 8.88; Score 80; DB 4; Length 539;
Best Local Similarity 28.88; Pred. No. 3;
Matches 17; Conservative 9; Mismatches 11; Indels 22; Gaps 3;

OY 17 CWLP-----WRTWWSSTAVSWASSALESTQPGATWT---KMLHYAG 61
Db 398 CWLSPGLRTASGALWMSGAWAAPSSW-----STSLRPGGSPWTGCGWLQVIG 448

RESULT 9

PCT-US95-09261-2
Sequence 2, Application PC/TUS9509261
GENERAL INFORMATION:
APPLICANT:
APPLICANT: NAME: BOARD OF REGENTS, THE UNIVERSITY OF TEXAS SYSTEM
APPLICANT: STREET: 201 West 7th Street
APPLICANT: CITY: Austin
APPLICANT: STATE: Texas
APPLICANT: COUNTRY: United States of America
APPLICANT: POSTAL CODE: 78701
APPLICANT: TELEPHONE NO: (512)499-4462
APPLICANT: TELEFAX: (512)499-4523
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE EXPRESSION OF
TITLE OF INVENTION: A BONE AND PROSTATE DERIVED GROWTH FACTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09261
FILING DATE: CONCURRENTLY HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/283,701


```

: FILING DATE: 01-AUG-1994
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: SERTICH, GARY J.
: REGISTRATION NUMBER: 34,430
: REFERENCE/DOCKET NUMBER: UTPC422P--
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (512) 418-3000
: TELEFAX: (713) 789-2679
: TELEEX: 79-0924
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 539 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US95-09261-2

Query Match      8.8%; Score 80; DB 5; Length 539;
Best Local Similarity 28.8%; Pred. No. 3;
Matches 17; Conservative 9; Mismatches 11; Indels 22; Gaps 3;

Cy 17 CMLP-----WRTWMSSSTAVMSASSALESTPATGATWT--KMLHYAG 61
Db 398 CMLSPGLRRTRASGALMWSGAMAAAPSSW-----STSLRPAGGPPWTGRCQWLQYLG 448

RESULT 10
US-08-652-816A-10
: Sequence 10, Application US/08652816A
: Patent No. 5872215
: GENERAL INFORMATION:
: APPLICANT: Osbourn, JK
: APPLICANT: Allen, DJ
: TITLE OF INVENTION: Specific binding members, materials and
: TITLE OF INVENTION: methods.
: NUMBER OF SEQUENCES: 53
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/652,816A
: FILING DATE: 23-MAY-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9125579.4
: FILING DATE: 02-DEC-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9125579.8
: FILING DATE: 02-DEC-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9206318.9
: FILING DATE: 24-MAR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9206372.6
: FILING DATE: 23-SEP-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9525004.9
: FILING DATE: 07-DEC-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9610824.6
: FILING DATE: 23-MAY-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/GB92/02240
: FILING DATE: 02-DEC-1992

```

```

: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/244,597
: FILING DATE: 01-JUN-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: David W. Clough
: REGISTRATION NUMBER: 36,107
: REFERENCE/DOCKET NUMBER: 28111/33308
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-474-6300
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 119 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: US-08-652-816A-10

Query Match      8.7%; Score 79.5; DB 2; Length 119;
Best Local Similarity 34.0%; Pred. No. 0.5;
Matches 17; Conservative 10; Mismatches 16; Indels 7; Gaps 2;

Cy 27 SSSSTAVMSASSALESTPATGATWTKMLHYAGSRISPTLENTLTVS 76
Db 28 SSSSYW-SWTRQ-----PPGKGLMIGYIHSGSTNSNSISRVITIS 70

RESULT 11
US-08-480-774A-2
: Sequence 2, Application US/08480774A
: Patent No. 5852186
: GENERAL INFORMATION:
: APPLICANT: MARASCO, Wayne A.
: APPLICANT: SODROSKI, Joseph G.
: APPLICANT: HASELTINE, William A.
: APPLICANT: POSNER, Marshall JR.
: TITLE OF INVENTION: REACTIVE NEUTRALIZING HUMAN
: TITLE OF INVENTION: ANTI-9P 120 RECOMBINANT ANTIBODY, DNA CODING THE SAME
: TITLE OF INVENTION: AND USE THEREOF
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/480,774A
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/400,674
: FILING DATE: 08-MAR-1995
: APPLICATION NUMBER: 07/804,652
: FILING DATE: 10-DEC-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Resnick, David S
: REGISTRATION NUMBER: 34,235
: REFERENCE/DOCKET NUMBER: 41450-FWC-DIV
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: TELEX:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 142 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear

```

MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-480-774A-2

Query Match 8.7%; Score 79.5; DB 2; Length 142;
Best Local Similarity 34.0%; Pred. No. 0.62;
Matches 17; Conservative 11; Mismatches 15; Indels 7; Gaps 2;

QY 27 SSSSTAWVSWASSALETSTQPATGATWTKMLHAGSSRISPTLEATLTVS 76
DB 47 SISSHYW-SWIRQ-----PCKGLQIGWIGYIYSGSTYNNPSLKSRYTIS 89

RESULT 12
US-08-545-809A-114
Sequence 114, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Matsuda, Fumihiko
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545, 809A
FILING DATE: 27-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29, 066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 114:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-545-809A-114

Query Match 8.7%; Score 79; DB 3; Length 117;
Best Local Similarity 29.4%; Pred. No. 0.55;
Matches 15; Conservative 13; Mismatches 17; Indels 6; Gaps 1;

QY 26 WSSSTAWVSWASSALETSTQPATGATWTKMLHAGSSRISPTLEATLTVS 76
DB 46 YSISSSNMWGWIRQ-----PCKGLEWIGYIYSGSTYNNPSLKSRYTIS 90

RESULT 13
US-09-260-527-3
Sequence 3, Application US/09260527A
Patent No. 6228599
GENERAL INFORMATION:
APPLICANT: Knox, J.P.
APPLICANT: Mikkelson, J.D.

APPLICANT: Willats, W. G.
TITLE OF INVENTION: ANTIBODY
FILE REFERENCE: DYO019.001AUS
CURRENT APPLICATION NUMBER: US/09/260,527A
CURRENT FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 278
TYPE: PRT
ORGANISM: UNKNOWN
FEATURE:
OTHER INFORMATION: Anti-homogalacturonan specific antibodies selected
OTHER INFORMATION: from a naive phage display library known as the
OTHER INFORMATION: Synthetic scfv library (#1) from the Centre for
OTHER INFORMATION: Protein Engineering, MRC Centre, Cambridge, UK
US-09-260-527-3

Query Match 8.7%; Score 79; DB 4; Length 278;
Best Local Similarity 29.4%; Pred. No. 1.6;
Matches 15; Conservative 13; Mismatches 17; Indels 6; Gaps 1;

QY 26 WSSSTAWVSWASSALETSTQPATGATWTKMLHAGSSRISPTLEATLTVS 76
DB 49 YSISSSNMWGWIRQ-----PCKGLEWIGYIYSGSTYNNPSLKSRYTIS 93

RESULT 14
US-08-851-362D-22
Sequence 22, Application US/08851362D
Patent No. 6235883
GENERAL INFORMATION:
APPLICANT: Jakobovits, Aya
APPLICANT: Yang, Xiao-Dong
APPLICANT: Gallo, Michael
APPLICANT: Jia, Xiao-Chi
TITLE OF INVENTION: Human Monoclonal Antibodies to Epidermal
TITLE OF INVENTION: Growth Factor Receptor
FILE REFERENCE: Cell 4.20
CURRENT APPLICATION NUMBER: US/08/851,362D
CURRENT FILING DATE: 1997-05-05
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22
LENGTH: 76
TYPE: PRT
ORGANISM: human
US-08-851-362D-22

Query Match 8.4%; Score 76.5; DB 4; Length 76;
Best Local Similarity 32.0%; Pred. No. 0.57;
Matches 16; Conservative 11; Mismatches 16; Indels 7; Gaps 2;

QY 27 SSSSTAWVSWASSALETSTQPATGATWTKMLHAGSSRISPTLEATLTVS 76
DB 7 SSGSYTW-SWIRQ-----PCKGLEWIGYIYSGSTYNNPSLKSRYTIS 49

RESULT 15
US-08-545-809A-142
Sequence 142, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Matsuda, Fumihiko
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA

```

1      COUNTRY: US
2      ZIP: 02110-2804
3      COMPUTER READABLE FORM:
4      MEDIUM TYPE: Diskette
5      COMPUTER: IBM Compatible
6      OPERATING SYSTEM: Windows95
7      SOFTWARE: FastSO for Windows Version
8      CURRENT APPLICATION DATA:
9      APPLICATION NUMBER: US/08/545,809A
10     FILING DATE: 27-MAR-1996
11     PRIOR APPLICATION DATA:
12     APPLICATION NUMBER: PC9/Jf93/00603
13     FILING DATE: 10-MAY-1993
14     ATTORNEY/AGENT INFORMATION:
15     NAME: Freeman, John W.
16     REGISTRATION NUMBER: 29,066
17     REFERENCE/DOCKET NUMBER: 06501/004001
18     TELECOMMUNICATION INFORMATION:
19     TELEPHONE: 617-542-5070
20     TELEFAX: 617-542-8906
21     TELEX: 200154
22     INFORMATION FOR SEQ ID NO: 142:
23     SEQUENCE CHARACTERISTICS:
24     LENGTH: 118 amino acids
25     TYPE: amino acid
26     TOPOLOGY: linear
27     MOLECULE TYPE: protein
28     US-08-545-809A-142

```

Query Match	8.4%	Score 76.5	DB 3	Length 118
Best Local Similarity	32.0%	Pred. NO. 1		
Matches	16	Conservative	11	Mismatches 16; Indels 7; Gaps 2.
QY	27	SSSSIAWYMASSALETSTQPTATGATWTKMLHYAASSRSPLLENTLYYS	76	
DB	49	SSGSIIYV-SWIRQ-----PQKGLGEMGIIYYLSTSTNPNPSTKRVTTS	91	

Search completed: March 27, 2003, 16:13:22
Job time : 36.0526 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 16:12:13 ; Search time 38.8421 Seconds
(without alignments)
244.921 Million cell updates/sec

Title: US-10-019-219-1

Perfect score: 912
Sequence: 1 TVVRLFLAMPCMWPCWLP.....MAACGARVKRRFLQTSLSR 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 237916 segs, 58723674 residues

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA.*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	10.2	145	US-09-984-271-240	Sequence 240, App
2	91	10.0	265	US-09-903-456-77	Sequence 77, App
3	88	9.6	1224	US-10-217-774-4	Sequence 4, Appl
4	87.5	9.6	890	US-10-060-425-8	Sequence 8, Appl
5	87.5	9.6	1422	US-08-424-550B-81	Sequence 81, Appl
6	85.5	9.4	890	US-10-060-425-10	Sequence 10, Appl
7	83	9.1	63	US-09-864-761-41103	Sequence 41103, A
8	81	8.9	1098	US-09-712-363-288	Sequence 288, App
9	80	8.8	162	US-09-738-626-6604	Sequence 6604, App
10	80	8.8	177	US-10-050-704-316	Sequence 316, App
11	79	8.7	90	US-09-864-761-48222	Sequence 48222, A
12	79	8.7	96	US-10-194-975-101	Sequence 101, App
13	79	8.7	98	US-10-194-975-35	Sequence 35, Appl
14	77.5	8.5	119	US-09-860-670-121	Sequence 121, Appl
15	77	8.4	150	US-10-038-010-4	Sequence 1674, Ap
16	76.5	8.4	638	US-10-038-010-4	Sequence 4, Appl
17	76.5	8.4	76	US-09-187-693-38	Sequence 38, Appl
18	76.5	8.4	99	US-10-194-975-43	Sequence 43, Appl
19	76.5	8.4	516	US-09-908-670-3	Sequence 3, Appl

20	76.5	8.4	890	9	US-10-060-425-2	Sequence 2, Appl
21	76	8.3	98	9	US-10-194-975-34	Sequence 34, Appl
22	76	8.3	209	10	US-09-904-536-8	Sequence 8, Appl
23	76	8.3	209	10	US-09-904-536-9	Sequence 9, Appl
24	76	8.3	209	10	US-09-904-536-11	Sequence 11, Appl
25	76	8.3	209	10	US-09-904-536-12	Sequence 12, Appl
26	76	8.3	209	10	US-09-904-536-13	Sequence 13, Appl
27	76	8.3	209	10	US-09-904-536-14	Sequence 14, Appl
28	76	8.3	209	10	US-09-904-536-15	Sequence 15, Appl
29	76	8.3	209	10	US-09-904-536-17	Sequence 17, Appl
30	76	8.3	209	10	US-09-904-536-18	Sequence 18, Appl
31	76	8.3	212	10	US-09-904-536-10	Sequence 10, Appl
32	76	8.3	235	9	US-10-095-449-6	Sequence 6, Appl
33	76	8.3	235	10	US-09-448-378-1	Sequence 1, Appl
34	76	8.3	235	10	US-09-983-806-6	Sequence 6, Appl
35	76	8.3	235	10	US-09-904-536-1	Sequence 1, Appl
36	76	8.3	635	9	US-10-099-895-33	Sequence 33, Appl
37	75.5	8.3	213	9	US-09-989-920-206	Sequence 206, App
38	75.5	8.3	213	9	US-09-989-920-251	Sequence 251, App
39	75.5	8.3	650	9	US-09-738-626-3856	Sequence 3856, Ap
40	75	8.2	138	10	US-09-867-550-1422	Sequence 1422, Ap
41	75	8.2	212	9	US-09-738-626-5989	Sequence 5989, Ap
42	74.5	8.2	1680	10	US-09-788-043C-5	Sequence 5, Appl
43	74.5	8.2	97	9	US-10-194-975-42	Sequence 42, Appl
44	74.5	8.2	97	9	US-10-194-975-52	Sequence 52, Appl
45	74.5	8.2	97	9	US-10-194-975-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1
US-09-984-271-240
Sequence 240, Application US/09984271
Publication No. US20030040088A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 71 Human Secreted Proteins
FILE REFERENCE: P2030P1
CURRENT APPLICATION NUMBER: US/09/984, 271
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: 09/482, 273
PRIOR FILING DATE: 2000-01-13
PRIOR APPLICATION NUMBER: PCT/US99/15849
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: 60/092, 921
PRIOR FILING DATE: 1998-07-15
PRIOR APPLICATION NUMBER: 60/092, 922
PRIOR FILING DATE: 1998-07-15
PRIOR APPLICATION NUMBER: 60/092, 956
PRIOR FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 267
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 240
LENGTH: 145
TYPE: PRT
ORGANISM: Homo sapiens
US-09-984-271-240
Query Match 10.2%; Score 93; DB 9; Length 145;
Best Local Similarity 26.5%; Pred. No. 0.18;
Matches 36; Conservative 11; Mismatches 61; Indels 28; Gaps 5;
QY 12 CMAYPCULPRTWMSSTRAWVW-----ASALETSTQPATGATWIKMLHYASSR 64
DB 12 CLWCPILPAMPPL--RGPRRSAMKRWPLPYGPAALGCSMTROTPTAVSWPCWL--MSSSL 67
QY 65 ISPTLEATVIVSPPLAARVAVCLRLCPYPKDSSTERSRVAMPSCPSLPAQLMSS 124
DB 68 STACLATLTIGSLARETRRR-----SLSPFWNSANQVPPSPPHSGIGR 113
QY 125 PRW--WPCLPVTKL 137

DB 114 RGMACHLTCCLVTL 129

RESULT 2
US-09-903-456-77

; Sequence 77, Application US/09903456
; Patent No. US20020138874A1

; GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories

; APPLICANT: Mukerji, Pradip

; APPLICANT: Leonard, Amanda Eun-Yeong

; APPLICANT: Huang, Yung-Sheng

; APPLICANT: Pereira, Suzette L.

; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF

; FILE REFERENCE: 6407.US.P3

; CURRENT APPLICATION NUMBER: US/09/903,456

; CURRENT FILING DATE: 2001-07-11

; PRIOR APPLICATION NUMBER: US 09/624,670

; PRIOR FILING DATE: 2000-07-24

; PRIOR APPLICATION NUMBER: US 09/379,095

; PRIOR FILING DATE: 1999-08-23

; PRIOR APPLICATION NUMBER: US 09/145,828

; PRIOR FILING DATE: 1998-09-02

; NUMBER OF SEQ ID NOS: 116

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 77

; LENGTH: 265

; TYPE: PRT

; ORGANISM: Thraustochytrium aureum

US-09-903-456-77

Query Match 10.0%; Score 91; DB 10; Length 265;

Best Local Similarity 24.4%; Pred. No. 0.52;

Matches 53; Conservative 16; Mismatches 60; Indels 88; Gaps 10;

DB 16 PCW-LPMTWMSST---STAVNS-----WASSA----- 40

DB 40 PCWHRMOTWPCSSASRSRSLSSRSPSSCTTCFSDPCPCAMFPAISSEAK 99

DB 41 -LET-----STOPATGATWTKMLHAGSSRISPTLEATLVSPFLA 80

DB 100 CLETTWRRATISMLRACLASCTCTCPRTSTSMIPSSFPARSSTRPSCKT-TMPPFLP 158

DB 81 SLRVARVCR-----LLCPY-----PKDSTERSWRVAMSCASLPAQLM 122

DB 139 SGLSPSTLQEVMTPTFQSSLTCTPSCHTTSSPPGSSGSSSRTSPFRSS----- 212

DB 123 SSPRMWPTCLPYTKTLRPMWMAACGARYRRPQLTSS 159

DB 213 ----WQCLSPCTTTSN-----ATTNRLICFLECTS 241

RESULT 3

US-10-217-774-4

; Sequence 4, Application US/10217774

; Publication No. US20020193583A1

; GENERAL INFORMATION:

; APPLICANT: Fiddle, Carl Johan

; APPLICANT: Hilbun, Erin

; TITLE OF INVENTION: No. US20020193583A1 Human Proteases and Polynucleotides Encodir

; FILE REFERENCE: LEX-0219-USA

; CURRENT APPLICATION NUMBER: US/10/217,774

; CURRENT FILING DATE: 2002-08-12

; PRIOR APPLICATION NUMBER: US/09/930,872

; PRIOR FILING DATE: 2001-08-14

; PRIOR APPLICATION NUMBER: US 60/225,852

; PRIOR FILING DATE: 2000-08-16

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 1224

; TYPE: PRT

; ORGANISM: homo sapiens
US-10-217-774-4

Query Match 9.6%; Score 88; DB 9; Length 1224;

Best Local Similarity 25.5%; Pred. No. 5.2;

Matches 37; Conservative 18; Mismatches 56; Indels 34; Gaps 7;

DB 12 CMNVPCWLPMTWMSSTTAVNSWASSALESTOPATGATWTKMLHAGSSRISPTLEA 71

DB 1042 CLTORCHKPKKIQMLVSS---AW-SQCVTCERGCTQKRFLLCAEKY--SGKYRELASKKC 1095

DB 72 TLTVSPFLASLRVARVCLILCPYPKDSTERSWRVAMSPSCASLPAQLMSPRWWPTC 131

DB 1096 SHLPKP---SLELRACADLPCCRHPPFAAGPS-RGSMFASP-----WSQC 1138

DB 132 LPVTKTLRPMWMAACGARYRRPQL 156

DB 1139 T-----ASCQGGVQTRSVQ 1152

RESULT 4

US-10-060-425-8

; Sequence 8, Application US/10060425

; Patent No. US20020164650A1

; GENERAL INFORMATION:

; APPLICANT: Hiebsch, Ronald

; TITLE OF INVENTION: Methods of Assessing Wolframin Protein Activity

; FILE REFERENCE: 00450.US1

; CURRENT APPLICATION NUMBER: US/10/060,425

; CURRENT FILING DATE: 2002-01-30

; PRIOR APPLICATION NUMBER: 60/266,385

; PRIOR FILING DATE: 2001-02-02

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 8

; LENGTH: 890

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-10-060-425-8

Query Match 9.6%; Score 87.5; DB 9; Length 890;

Best Local Similarity 21.2%; Pred. No. 4;

Matches 28; Conservative 20; Mismatches 45; Indels 39; Gaps 5;

DB 1 TYVRLFLANLPCMWPCW-----LPWRTWMSSTTAVN--VSWASSALET 43

DB 633 SYIKLILVWLTAIILFCMEFYVRRSEGMKYNSTLTWQGVGFCGPRAKMETNARTQILC 692

DB 44 STOPATGATWTKMLHAGSSRISPTLEATLVSPFLAS--LVARVCLRLCPYPKSS 101

DB 693 SHLEGHRYVTGTFYRYVTEIDNSAESAINMLPFLGDMR-----CL----- 736

DB 102 TEPSWRVAPSC 113

DB 737 ----YGEATPSC 744

RESULT 5

US-08-424-550B-81

; Sequence 81, Application US/08424550B

; Patent No. US20020119447A1

; GENERAL INFORMATION:

; APPLICANT: JOHN N. SIMONS

; APPLICANT: TAMU J. PILOT-MATIAS

; APPLICANT: GEORGE J. DAMSON

; APPLICANT: GEORGE G. SCHLAUDER

; APPLICANT: SURESH M. DESAI

; APPLICANT: THOMAS P. LEARY

; APPLICANT: ANTHONY SCOTT MUEHROFF

; APPLICANT: JAMES C. ERKER

; APPLICANT: SHERI L. BUTIK

; APPLICANT: ISA K. MUSHAMMAR

; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS


```

: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC022045.2
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
: OTHER INFORMATION: EST_HUMAN HIT: BE008547.1, EVALU6.00e-15
US-09-864-761-41103

Query Match
Best Local Similarity 9.1%; Score 83; DB 10; Length 63;
Matches 24; Conservative 3; Mismatches 25; Indels 10; Gaps 3;

OY 93 CDPYKDSSTPS-----WVAM-PCSPASLPQALMSRPMPTCLPVTKLTRPMW 143
DB 3 CRYATPSSCMWSCRYPACHPWLRAWPPSPSWSASPSOP-TPTCSSAASPSLRSSW 61
OY 144 AA 145
DB 62 AS 63

RESULT 8
US-09-712-363-288
: Sequence 288, Application US/09712363
: Patent No. US20020164588A1
: GENERAL INFORMATION:
: APPLICANT: Eisenberg, David
: APPLICANT: Rotstein, Sergio H.
: APPLICANT: Marcotte, Edward M.
: TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
: FILE REFERENCE: 07419-032001
: CURRENT APPLICATION NUMBER: US/09/712,363
: CURRENT FILING DATE: 2000-11-13
: PRIOR APPLICATION NUMBER: PCT/US00/02246
: PRIOR FILING DATE: 2000-01-28
: PRIOR APPLICATION NUMBER: 60/179,531
: PRIOR FILING DATE: 2000-02-01
: PRIOR APPLICATION NUMBER: 60/117,844
: PRIOR FILING DATE: 1999-01-29
: PRIOR APPLICATION NUMBER: 60/118,206,
: PRIOR FILING DATE: 1999-02-01
: PRIOR APPLICATION NUMBER: 60/126,593
: PRIOR FILING DATE: 1999-03-26
: PRIOR APPLICATION NUMBER: 60/134,093
: PRIOR FILING DATE: 1999-05-14
: PRIOR APPLICATION NUMBER: 60/134,092
: PRIOR FILING DATE: 1999-05-14
: PRIOR APPLICATION NUMBER: 60/165,124
: PRIOR FILING DATE: 1999-11-12
: PRIOR APPLICATION NUMBER: 60/165,086
: PRIOR FILING DATE: 1999-11-12
: NUMBER OF SEQ ID NOS: 292
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 288
: LENGTH: 1098
: TYPE: PRT
: ORGANISM: Mycobacterium tuberculosis
US-09-712-363-288

Query Match
Best Local Similarity 8.9%; Score 81; DB 9; Length 1098;
Matches 29; Conservative 16; Mismatches 50; Indels 14; Gaps 3;

OY 6 FLAWLPCMMVPCMLPWRTWMMSSSTAWSSALETSTOP-----ATGATWTKMLH 58
DB 628 FLAALFLALCWAATTNOMWTVSYSGVFPNSAMPKIDGITVSTIFFALFAIAGYAAWMLH 687
```

```

OY 59 YAGSSNISPTLENTLVSP-----FLASLRVAVCLRLCPYPYKDS 101
DB 688 FAPRGAGEGRLLRALTAPVIVAGFMAAVFVAMAGIV-RQYPTYSN 735

RESULT 9
US-09-738-626-6604
: Sequence 6604, Application US/09738626
: Publication No. US20020197605A1
: GENERAL INFORMATION:
: APPLICANT: NAKAGAWA, SATOSHI
: APPLICANT: MIZOGUCHI, HIROSHI
: APPLICANT: ANDO, SEIKO
: APPLICANT: HAYASHI, MIKIRO
: APPLICANT: OCHIAI, KEIKO
: APPLICANT: YOKOI, HARUHIKO
: APPLICANT: TATEISHI, NAOKO
: APPLICANT: SENOH, AKIHIRO
: APPLICANT: IKEDA, MASATO
: APPLICANT: OZAKI, AKIO
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-125
: CURRENT APPLICATION NUMBER: US/09/738,626
: CURRENT FILING DATE: 2000-12-18
: PRIOR APPLICATION NUMBER: JP 99/377484
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: JP 00/159162
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: JP 00/280988
: PRIOR FILING DATE: 2000-08-03
: NUMBER OF SEQ ID NOS: 7059
: SOFTWARE: PatentIn ver. 3.0
: SEQ ID NO 6604
: LENGTH: 162
: TYPE: PRT
: ORGANISM: Corynebacterium glutamicum
US-09-738-626-6604

Query Match
Best Local Similarity 8.8%; Score 80; DB 9; Length 162;
Matches 44; Conservative 17; Mismatches 55; Indels 36; Gaps 9;

OY 18 WLPWRWMMSSSTAWSSALETSTOPATGA--TTWKLHYAGSSNISPTLENTLV 75
DB 6 WLPORVPSAVSSPSSSHQPNRCPHSAAPPSTSHW-----HSRSP----- 54
OY 76 SPFLASLRVAVCLRLCPYPYKDSSTPSWVAMPSCPS-----LPAQLMSSP 125
DB 55 PWCPL-----PCLNDRPSF--SASALRPSASIRSPASSASNSGHRSTFEPSLLPSR 108
OY 126 RW-WPTCLPVTRLTPRPMWMAAGARYKRRFLQ 156
DB 109 FWHMEK-----FTLQPAHAYV-KQIWRRLQ 134

RESULT 10
US-10-050-704-316
: Sequence 316, Application US/10050704
: Publication No. US20030050442A1
: GENERAL INFORMATION:
: APPLICANT: Ruben et al.
: TITLE OF INVENTION: 62 Human Secreted Proteins
: FILE REFERENCE: P2039P1
: CURRENT APPLICATION NUMBER: US/10/050,704
: CURRENT FILING DATE: 2002-01-18
: PRIOR APPLICATION NUMBER: 09/684,524
: PRIOR FILING DATE: 2000-10-10
: PRIOR APPLICATION NUMBER: PCT/US00/08979
: PRIOR FILING DATE: 2000-04-06
: PRIOR APPLICATION NUMBER: 60/128,693
: PRIOR FILING DATE: 1999-04-09
: PRIOR APPLICATION NUMBER: 60/130,991
```



```

; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 316
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-050-704-316

Query Match
Best Local Similarity 22.1%; Pred. No. 3.3;
Matches 33; Conservative 18; Mismatches 52; Indels 46; Gaps 7;

QY 44 STOPGATGAT-----TKMLHAGSSSRISPTLEATLTVSPF-----LASLRVA-----RVC 88
D 1 SASCATGSSSRVGTGLTLPDRHSSQAAPTLRASMWLATFRWIMPGICTTLELWAMPISLRGC 60

QY 89 LRLCPYPYRDSSTEPSRVRWAMPSC-----PASLPQOLMSSPRMWP--- 130
D 61 SLFACVLTSTKPPPRMTWLMPICEYCWLSPLSPSSVSSLSMRVLSGNMGPCVSG 120

QY 131 CL-PVTKLTLP-----WNAAC 146
D 121 CVSSISSEMAPSATISIGSPQTHWMLHC 149

RESULT 11
US-09-864-761-48222
; Sequence 48222, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
```

```

; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48222
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AB019441.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.42
; OTHER INFORMATION: EST_HUMAN HIT: BE295500.1, EVALUOE 2.00e-07
; OTHER INFORMATION: SWISSPROT HIT: O93571, EVALUOE 2.90e+00
US-09-864-761-48222

Query Match
Best Local Similarity 25.2%; Pred. No. 1.9;
Matches 30; Conservative 9; Mismatches 40; Indels 40; Gaps 7;

QY 15 VPC-WLPMRTWMMSSSTAWVSNASALETSTOPATGATWTKMLHAGSSSRISPTLEATL 73
D 4 VPARWMPVHRKMW-----LNRNALPPCANAS--LSSANAWLISCF-----SLEFTD 49

QY 74 T-----VSPFLASLRVARVCLRLCPYPYRDSSTEPSRVRWAMPSCPASLPA-OLMSSPR 126
D 50 CAIIGGIQPCACSTSGRSCQHV-----PACPATQACACNLGSDPR 89

RESULT 12
US-10-194-975-101
; Sequence 101, Application US/10194975
; Publication No. US20030039649A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01
; CURRENT APPLICATION NUMBER: US/10/194,975
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 101
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-194-975-101

Query Match
Best Local Similarity 34.0%; Pred. No. 2.1;
Matches 17; Conservative 11; Mismatches 14; Indels 8; Gaps 2;

QY 27 SSSSTANVSNASALETSTOPATGATWTKMLHAGSSSRISPTLEATLTVS 76
D 28 SISSYYW-SWI-----RQPPKLEWIGYIYSGSYNPNLSKSRVTIS 69

RESULT 13
US-10-194-975-35
; Sequence 35, Application US/10194975
; Publication No. US20030039649A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01
; CURRENT APPLICATION NUMBER: US/10/194,975
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Patentin version 3.1
```

SEQ ID NO 35
LENGTH: 98
TYPE: PRT
ORGANISM: Homo sapiens
US-10-194-975-35

Query Match 8.7%; Score 79; DB 9; Length 98;
Best Local Similarity 29.4%; Pred. No. 2.1;
Matches 15; Conservative 13; Mismatches 17; Indels 6; Gaps 1;

QY 26 WSSSTAWWSAASLETSTOPATGATWTKLHYAGSSRISPTLEATLTVS 76
DB 27 YSISSNMWGMIRQ-----PPGKGLWIGIYISGSIYTNESLSKSRVIMS 71

RESULT 14
US-09-860-670-121
Sequence 121, Application US/09860670
Patent No. US20020165137A1
GENERAL INFORMATION:

APPLICANT: Ruben et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA127P1
CURRENT APPLICATION NUMBER: US/09/860,670
CURRENT FILING DATE: 2001-05-21
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 289
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 121
LENGTH: 119
TYPE: PRT
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: SITE
LOCATION: (102)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (119)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-860-670-121

Query Match 8.5%; Score 77.5; DB 9; Length 119;
Best Local Similarity 31.2%; Pred. No. 3.6;
Matches 30; Conservative 9; Mismatches 34; Indels 23; Gaps 6;

QY 51 ATWTKLHYAGSSRISPT-LEATLTVSPFLASLRYAVCLRL-----CPYPKDS 101
DB 4 ATGRKWLWESGATGSSATPLPPLTSCPLCP1----PCSHLVLTSDGSGCSPPPQSGA 59
QY 102 -----TSPSRVAMP-----CPASLPALMSSPRW 128
DB 60 GVQAPRTSPWESLPSHCHLC-ASLGSPISEAPHLW 94

RESULT 15
US-09-867-550-1674
Sequence 1674, Application US/09867550
Patent No. US20020082206A1
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.
APPLICANT: Mehraban, Fued,
APPLICANT: Conley, Pamela
APPLICANT: Law, Debbie
APPLICANT: Topper, James
TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
TITLE OF INVENTION: Thereby
FILE REFERENCE: 21402-013 (Cura-313)
CURRENT APPLICATION NUMBER: US/09/867,550
CURRENT FILING DATE: 2001-09-20
PRIOR APPLICATION NUMBER: USSN 60/208,427
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 2125
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1674
LENGTH: 150
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (55)
OTHER INFORMATION: wherein Xaa may be any one of Ala or Arg or Gln or Glu or Gly
OTHER INFORMATION: Lys or Met or Pro or Ser or Thr or Trp or Val
US-09-867-550-1674

Query Match 8.4%; Score 77; DB 10; Length 150;
Best Local Similarity 29.9%; Pred. No. 5.1;
Matches 43; Conservative 13; Mismatches 38; Indels 50; Gaps 10;

QY 26 WSSSTAWWSAASLETSTOPATGA---TWTKLHYAGSSRISPTLEATLTVSPFLASL 82
DB 15 WSEATP-----APSAVSSCRILAAARAFSTKALNFA-----VSGLXAR 58
QY 83 RVAR-----VCLRL-----LCPPYPKDSSTEPSWR---VAMPSCPALPAQL----- 121
DB 59 RGRSLPGRSSCPRLGTAAVCTP-----DPPWRGGAALRTGPGPLPDQRHRRPPPP 111
QY 122 --MSPRRMPTCLPYTKLTLRFW 143
DB 112 PGISPLPPVAMP-SGLSSASW 134

Search completed: March 27, 2003, 16:20:25
Job time : 39.8421 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 16:01:27 ; Search time 27.4737 Seconds

(without alignments)
566.862 Million cell updates/sec

Title: US-10-019-219-1

Perfect score: 912

Sequence: 1 TVVRLFLAMPCMMVPCWLP.....MAACGARVKRRFLQTLISLR 162

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	10.5	157	2 S73704	hypothetical prote
2	87.5	9.6	790	2 T34293	hypothetical prote
3	85.5	9.4	261	2 J00137	hypothetical 30.1k
4	84.5	9.3	210	2 H72460	hypothetical prote
5	83.5	9.2	384	2 G82670	hypothetical prote
6	83.5	9.2	436	2 A54770	general secretory
7	83	9.1	332	2 C72310	N-acetylglucosamin
8	83	9.1	377	2 F95948	conserved hypochet
9	81.5	8.9	254	2 B83286	probable acyltrans
10	81	8.9	984	2 T00326	hypothetical prote
11	81	8.9	1096	2 G70697	hypothetical prote
12	81	8.9	1522	2 T00028	probable arabinosy
13	80	8.8	610	2 S35049	brain-specific ang
14	79.5	8.7	660	2 J00067	mucin JER57 - huma
15	79	8.7	98	2 S26903	chitinase (EC 3.2.
16	79	8.7	98	2 S12414	Ig heavy chain V r
17	79	8.7	1007	2 T01437	Ig heavy chain V r
18	78.5	8.6	784	2 T00317	hypothetical prote
19	78	8.6	3570	2 T45025	hypothetical 82k p
20	77.5	8.5	253	2 AC0810	mucin MUC5B, trach
21	77.5	8.5	263	1 S43189	probable sulfate t
22	77	8.4	118	2 A72474	hypothetical prote
23	77	8.4	145	2 S78055	hypothetical prote
24	77	8.4	146	2 S09711	Ig heavy chain pre
25	77	8.4	211	2 S73791	Ig heavy chain V r
26	77	8.4	1513	2 A54895	hypothetical prote
27	77	8.4	1859	1 A34092	mucin 2, intestinal
28	77	8.4	1862	2 T29959	DNA-directed RNA p
29	76.5	8.4	99	2 S12412	Ig heavy chain V r

30	76.5	8.4	118	2 A26340	Ig heavy chain pre
31	76.5	8.4	130	2 S31690	Ig heavy chain V r
32	76.5	8.4	290	1 BVEBC2	Cys2 protein - Sal
33	76.5	8.4	304	2 S39892	nfd protein - Rho
34	76.5	8.4	466	2 AD3477	malate dehydrogena
35	76.5	8.4	516	2 S19586	N-methyl-D-asparta
36	76.5	8.4	580	2 T43481	probable mucin DKF
37	76.5	8.4	825	2 S75173	hypothetical prote
38	76.5	8.4	947	2 B97567	valyl-tRNA synthet
39	76.5	8.4	947	2 A12787	valyl-tRNA synthet
40	76.5	8.4	1711	1 A55148	protein-tyrosine-P
41	76	8.3	98	2 S26905	Ig heavy chain V r
42	76	8.3	157	2 S09805	hypothetical prote
43	76	8.3	198	2 T35915	hypothetical prote
44	76	8.3	235	2 T38440	ftt3 ligand - huma
45	76	8.3	296	2 JC7268	paired-type homeod

ALIGNMENTS

RESULT 1

S73704 hypothetical protein H08_orf157a - Mycoplasma pneumoniae (strain ATCC 29342)

C:Species: Mycoplasma pneumoniae

A:Variety: ATCC 29342

C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 15-Sep-2000

C:Accession: S73704

R:Himmelreich, R.; Hilbert, H.; Piagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.

Nucleic Acids Res. 24, 4420-4449, 1996

A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumon

A:Reference number: S73327; MUID:97105885; PMID:8948633

A:Accession: S73704

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-157 <HIM>

A:Cross-references: EMBL:AE000036; GB:U00089; NID:91674053; PIDN:AA96026.1; PID:9167

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996

C:Genetics:

A:Genetic code: SGC3

C:Superfamily: Mycoplasma pneumoniae hypothetical protein H08_orf157a

Query Match 10.5%; Score 96; DB 2; Length 157;

Best Local Similarity 23.8%; Pred. No. 0.095;

Matches 36; Conservative 19; Mismatches 62; Indels 34; Gaps 8;

QY 9 WLPCCMVPVPCWLP-----W-RTWWSSTTAVSW-----ASSALETSTQ----- 46

Db 4 WMSC-APPITYPHNTSWTESGMDRTSMWRWSAQRWGSFKRIVRANKLRVMAKTKMPLY 62

QY 47 --PATGATWTKW-----LHYAGSSRSIPTLEATITVSPFLASLRVARVCRLICPPPK 98

Db 63 LIPSPNPKPYSKLAINGELITLPKKITSPPATSSLSIKRPPGRCYINARLSMR--CPTLSR 120

QY 99 DSTEPSWRVAMPSCPASLPQQLMSSPPRMP 129

Db 121 KVRV-PTIKVPMVRAPSTPKSTSSNNPMP 150

RESULT 2

T34293 hypothetical protein F49E10.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000

C:Accession: T34293

R:Miller, N.

submitted to the EMBL Data Library, April 1996

A:Description: The sequence of C. elegans cosmid F49E10.

A:Reference number: 221500

A:Accession: T34293

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-790 <MLT>

C/Accession: A54770: S37024
E:Blanck, O.; Perrin, C.; Mzaut, H.; Darbon, H.; Mattel, M.G.; Mjauellis, R.

Genomics 21, 18-26, 1994
A:Title: Molecular cloning, cDNA analysis, and localization of a monomer of the N-acetyl
A:Reference number: A54770; MID:94375011; PMID:8088785
A:Accession: A54770
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-46 <BLA>
A:Cross-references: EMBL:X72018
C:Genetics:
A:Gene: GDB:NAGR1
A:Cross-references: GDB:250465; OMIM:160994
A:Map position: 19p13.2-19p13.2
C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein
C:Keywords: calcium binding; lectin; thyroid gland; transmembrane protein
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-46/Product: N-acetylglucosamine receptor 1 status predicted <MAM>
F:45-111/Domain: ribonucleoprotein repeat homology <RMM2>
F:230-236/Region: glycine-rich
F:441-460/Domain: transmembrane #status predicted <TMM>
F:488-495/Region: coated-pit mediated internalization signal

Query Match 9.2%; Score 83.5; DB 2; Length 496;
Best Local Similarity 23.7%; Pred. No. 4.2;
Matches 27; Conservative 13; Mismatches 35; Indels 39; Gaps 6;

QY 14 MYPCLMRTWSSSSSTAVWSSASLETSTPATGATWTKMLHYGSSRISPLENTL 73
Db 391 MPAMSAWAPTIW--SGWAMSAWAPTA-----SSAMWAMWAPPTASSAWAP----- 435
QY 74 TVSFSLARVACRLCLCPYKDSSTPSRVAMPSCASLPALQMSPPRW 127
Db 436 WAPWALAL-----SAW--AMPWAVAVPALTPS-RW 465

RESULT 7
C72310
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: C72310
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72300; MUID:99287316; PMID:10360571
A:Accession: C72310
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-332 <ARN>
A:Cross-references: GB:AE001760; GB:AE000512; NID:94981510; PIDN:ADJ36061.1; PID:9498152
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0982

Query Match 9.1%; Score 83; DB 2; Length 332;
Best Local Similarity 26.8%; Pred. No. 3.1;
Matches 30; Conservative 20; Mismatches 42; Indels 20; Gaps 6;

QY 30 STAVWSSASLET-STOPATGATWTKMLH-----YAGSSRISPLENTLVS 76
Db 104 TTAFMALAYALTAHATKGLFSGWIKWLSNTYVTSNTNPTYA--PTGPTIATVLSIS 161
QY 77 PELSLARVACRLCLCPYKDSSTPSRVAMPSCASLPALQMSPPRW 128
Db 162 PWMS--ITFLALMLWTFGVAKNSQPS-KLNM--TASVLAITLAPLAW 208

RESULT 8
F95948
probable acyltransferase, possibly surface-saccharide specific acetyltransferase protein
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C:Accession: F95948
R:Finan, T.M.; Weldner, S.; Wong, K.; Buhrmester, J.; Chalm, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSyMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: F95948
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-377 <KUR>

A:Cross-references: GB:AL591985; PIDN:CAC49254.1; PID:915140740; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid PSyMB
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hub
pela, D.; Chalm, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
hebut, P.; Vandenbol, M.; Vorholter, F.J.; Weldner, S.; Wells, D.H.; Wong, K.; Yeh,
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMD21188
A:Genome: plasmid

Query Match 9.1%; Score 83; DB 2; Length 377;
Best Local Similarity 22.8%; Pred. No. 3.5;
Matches 36; Conservative 11; Mismatches 57; Indels 54; Gaps 7;

QY 18 WLPW---RTWMMSSSTAVWSSASLETSTPATGAT-----WTKMLHYAG 61
Db 120 WTPWYLCHTWSLSTIEQFLWPLVLVLPKRSVAGCVGVYCSLAVRTW----- 171
QY 62 SSRSISPLEATLVSPFLASLVARVCLRLCPYKDS-----STPSRVAMPSCP 114
Db 172 -----PLTGPPSLA-----RDLPLPASMALAVGALLARPSWRSRGPMA 212
QY 115 --ASLPAQIMSSPRMWPCTCLEPTKTLRPMWAAAGARY 150
Db 213 KLSWPLSLASLCLVWSK--PVAMPVAVWPAWIGLEV 248

RESULT 9
B83286
hypothetical protein PA2884 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83286

R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pa
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: B83286
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-254 <STO>
A:Cross-references: GB:AE004714; GB:AE004091; NID:9948965; PIDN:AA06272.1; GSPDB:GN
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA2884

Query Match 8.9%; Score 81.5; DB 2; Length 254;
Best Local Similarity 23.1%; Pred. No. 3.2;
Matches 37; Conservative 14; Mismatches 62; Indels 47; Gaps 7;

QY 2 VVRLFLAMPCMMVPCWLPWRTWSSSTAVWSSASL-----ET 43
Db 30 IYLYLIWV--TYAERWLPPTLLW-----PVSLAASALIAHYLPGRIPRLPMDLA 80
QY 44 STQPATGATWTKMLHYAGSSRISPLENTLVSPF--LASLAVAV--CLRLCLCPYPRD 99
Db 81 SASPPIQMTLWLPW---DKALVALTLMLWLRPRKOPLVSDITLALFCLTFVVPPLSI 136

OY 100 SSTEP5RVAMPSCPASLPAQLMSSPRWPTCLPYTKL 139
DB 137 MTNMA5WOPKMPDA-----FMMGLVLANGVSL 164

RESULT 10

T00326
hypothetical protein KIAA0550 - human
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00326
R:Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.
DNA Res. 5, 31-39, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. IX. The complet
A:Reference number: Z14086; MUID:98290545; PMID:9628581
A:Accession: T00326
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-984 <NAG>
A:Cross-references: EMBL:AB011122; NID:g3043623; PIDN:BAA25476.1; PID:g3043624
A:Experimental source: brain
C:Genetics:
A:Note: KIAA0550
C:Superfamily: thrombospondin type 1 repeat homology
F:344-398/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 8.9%; Score 81; DB 2; Length 984;
Best Local Similarity 21.6%; Pred. No. 14;
Matches 32; Conservative 13; Mismatches 59; Indels 44; Gaps 6;

OY 11 PCMNVP-----WLPRTWMS-----SSRAWVSSSALETST 45
DB 391 PCNIALCPVDGQWQEMSSQCSVTCSNCTQQRCTAAHGGSCRCRPMESRECIYP 450
OY 46 OPATGATWTKMLHYAGSS-----RISPTLEATLVSPFLASLRVAVCLRLCP-P 95
DB 451 ECTANGQNMWGMHWSGSCSCDGGMERIRITCOGAVITTOQCEGTEEVRRCSQRCRAP 510
OY 96 Y---PRDSSTEPSWR-----VAMPSCP 114
DB 511 YEICPEDYLM5WVKRTPAGDLAFNCP 538

RESULT 11

G70697
probable arabinosyltransferase - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: G70697
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, R.; Gentile, S.; Hamlin, N.; Holroyd, S.
; Rajandram, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:9629587; PMID:9634230
A:Accession: G70697
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1098 <COL>
A:Cross-references: GB:Z80343; GB:AL123456; NID:g3261648; PIDN:CAB02474.1; PID:g1552878
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: emdB

Query Match 8.9%; Score 81; DB 2; Length 1098;
Best Local Similarity 26.6%; Pred. No. 16;
Matches 29; Conservative 16; Mismatches 50; Indels 14; Gaps 3;

OY 6 FLAWLPCMWVPCMLPWRTWSSSTAWVSSALETSTOP-----ATGATWTKWLH 58
DB 628 FLAALFLLALCOWATTNCGWVYSSYGVPNSAMPKIDGITVSTIFFALFAIAGTAAMLH 687

OY 59 YAGSSRISPTLEATLVSP-----FLASLRVAVCLRLCPYPKDS 101
DB 688 FAPRGAGEGRLLIALTAPPIVAGFMAVFAVSMAGIV-RQYPYSN 735

RESULT 12

T00028
brain-specific angiogenesis inhibitor 3 - human
N:Alternate names: BA13 protein
C:Species: Homo sapiens (man)
C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C:Accession: T00028
R:Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.
Cytogenet. Cell Genet. 79, 103-108, 1997
A:Title: Cloning and characterization of BA12 and BA13, novel genes homologous to bra
A:Reference number: Z14066; MUID:98194217; PMID:9533023
A:Accession: T00028
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1522 <SHI>
A:Cross-references: EMBL:AB005299; NID:g3021700; PIDN:BAA25363.1; PID:g3021701
A:Experimental source: brain
C:Genetics:
A:Gene: GDB:BA13
A:Cross-references: GDB:9838090; OMIM:602684
A:Map position: 6q12-6q12
C:Superfamily: thrombospondin type 1 repeat homology
F:344-398/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 8.9%; Score 81; DB 2; Length 1522;
Best Local Similarity 21.6%; Pred. No. 22;
Matches 32; Conservative 13; Mismatches 59; Indels 44; Gaps 6;

OY 11 PCMNVP-----WLPRTWMS-----SSRAWVSSSALETST 45
DB 391 PCNIALCPVDGQWQEMSSQCSVTCSNCTQQRCTAAHGGSCRCRPMESRECIYP 450
OY 46 OPATGATWTKMLHYAGSS-----RISPTLEATLVSPFLASLRVAVCLRLCP-P 95
DB 451 ECTANGQNMWGMHWSGSCSCDGGMERIRITCOGAVITTOQCEGTEEVRRCSQRCRAP 510
OY 96 Y---PRDSSTEPSWR-----VAMPSCP 114
DB 511 YEICPEDYLM5WVKRTPAGDLAFNCP 538

RESULT 13

S35049
nucin JER57 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 10-Dec-1993 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: S35049; S37594
R:Dufosse, J.; Porchet, N.; Audie, J.P.; Guyonnet Duperrat, V.; Laine, A.; van-Seuning
Biochem. J. 293, 329-337, 1993
A:Title: Degenerate 87-base-pair tandem repeats create hydrophilic/hydrophobic altern
A:Reference number: S35047; MUID:93343858; PMID:7916618
A:Accession: S35049
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-610 <DUF>
A:Submitted to the EMBL Data Library, September 1993
A:Reference number: S37594
A:Accession: S37594
A:Molecule type: mRNA
A:Residues: 1-20, 'W', 22-610 <AUB>
A:Cross-references: EMBL:X74955

Query Match 8.8%; Score 80; DB 2; Length 610;
Best Local Similarity 23.9%; Pred. No. 11;
Matches 16; Conservative 16; Mismatches 19; Indels 16; Gaps 2;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 16:00:22 ; Search time 27.4737 Seconds

(without alignments)
244.567 Million cell updates/sec

Title: US-10-019-219-1

Perfect score: 912
Sequence: 1 TVVRLFLANLPCMMVPCWLP.....MAACGARVKRRRLQTLSTLSR 162

Scoring table: BLUOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	96	10.5	157	1	YE63_MYCPN
2	85.5	9.4	890	1	WFS1_MOUSE
3	81	8.9	1098	1	EMBB_MYCTU
4	81	8.9	1522	1	BA13_HUMAN
5	80.5	8.8	1595	1	LTBL_HUMAN
6	78.5	8.6	784	1	YAV2_XANCV
7	78.5	8.6	1065	1	EMBB_MYCAV
8	77.5	8.6	1139	1	M2A2_HUMAN
9	77.5	8.5	253	1	CYS2_SALTI
10	77	8.4	211	1	YD71_MYCPN
11	77	8.4	487	1	HEP_DROME
12	77	8.4	1852	1	RPB1_CAREL
13	76.5	8.4	253	1	CYS2_SALTY
14	76.5	8.4	890	1	WFS1_HUMAN
15	76.5	8.4	1711	1	PTPO_RAT
16	76	8.3	157	1	UL42_HCMVA
17	76	8.3	235	1	FL3L_HUMAN
18	76	8.3	635	1	TPOR_HUMAN
19	75.5	8.3	2845	1	APC_MOUSE
20	75	8.2	725	1	AGAL_YEAST
21	74.5	8.2	546	1	AAAS_HUMAN
22	74.5	8.2	951	1	HEX_ADE05
23	74.5	8.2	967	1	HEX_ADE02
24	74	8.1	348	1	HFE_HUMAN
25	74	8.1	1776	1	POLR_OYMY
26	73.5	8.1	108	1	YNEA_SALTY
27	72.5	7.9	138	1	YV91_MYCPN
28	72.5	7.9	540	1	YV91_MOUSE
29	72	7.9	108	1	YV91_MOUSE
30	72	7.9	498	1	YV91_MOUSE
31	72	7.9	715	1	YV91_MOUSE
32	72	7.9	1229	1	YV91_MOUSE
33	71.5	7.8	550	1	SOA1_CERAE

34	71.5	7.8	550	1	SOA1_HUMAN	P35610	homo sapien
35	71.5	7.8	550	1	SOA1_MACFA	O77761	macaca fasc
36	71.5	7.8	939	1	HEX_ADEP3	O9Y188	porcine ade
37	71	7.8	108	1	YNEA_ECO57	O8X766	escherichia
38	71	7.8	215	1	Y215_ADE02	P03227	human adeno
39	71	7.8	344	1	GUN4_TTRRE	O14405	trichoderma
40	71	7.8	437	1	EF1G_HUMAN	P26641	homo sapien
41	71	7.8	437	1	EF1G_RABIT	P29694	oryctolagus
42	71	7.8	484	1	PER2_VOLCA	P81133	volvox cart
43	71	7.8	836	1	GCSR_HUMAN	O99062	homo sapien
44	70.5	7.7	387	1	OR19_DROME	O91816	drosophila
45	70.5	7.7	556	1	CD19_HUMAN	P15391	homo sapien

ALIGNMENTS

RESULT 1	ID	YE63_MYCPN	STANDARD	PRT	157 AA
AC	P75320	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DE	Hypothetical protein MPN463 (H08_Ori157a).				
GN	MPN463 OR MP378.				
OS	Mycoplasma pneumoniae.				
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.				
NCBI_TaxID=2104;					
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 29342 / M129;				
RX	MEDLINE-97105885; PubMed-8948633;				
RA	Himmelfreich R., Hilbert H., Piagens H., Pirkl E., Li B.-C.,				
RA	Herrmann R.,				
RT	"Complete sequence analysis of the genome of the bacterium Mycoplasma				
RL	pneumoniae."				
CC	Nucleic Acids Res. 24:4420-4449(1996).				
CC	-1 SIMILARITY: ALMOST IDENTICAL TO M.PNEUMONIAE MPN091 AND MPN413.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL; AE000036; AAB96026.1; "				
RW	Hypothetical protein; Complete proteome.				
SO	SEQUENCE 157 AA; 17842 MW; B7F5CCB85395DB7 CNC64;				
Query Match	10.5%; Score 96; DB 1; Length 157;				
Best Local Similarity	23.8%; Pred. No. 0.054;				
Matches	36; Conservative 19; Mismatches 62; Indels 34; Gaps 8;				
OY	9 WPCMMVPCWLP-----W-RTWMMSSSTAVVSN-----ASSALESTO-----46				
DB	4 WMSC-APLITVDHNTSWSGMDRTSMRWMSQKRWGSEFKIVRANKALRYMAKTKMPLV 62				
OY	47 --PATGATWTKW-----LHYAGSSRIPTLEATLVSPFLASLVARVAVCLLCPYPK 98				
DB	63 LIPPEPNPKYSKLAINOELHLPHKKTSPATSSSLKPRGRCGYLNAHLNWR--CPTLSR 120				
OY	99 DSTSESRVAMPSCPASLPAQIMSSPRMP 129				
DB	121 KVRV-PTIKVPVVRAPSTKPSKTSSTSSNNPMP 150				
RESULT 2					
WFS1_MOUSE					
ID	WFS1_MOUSE	STANDARD	PRT	890 AA	
AC	P56695; Q92276;				

[illegible]

DE DE

DE 15-JUN-2002 (Rel. 4), Last annotation update)

DE latent transforming growth factor beta binding protein 1L precursor

DE (Transforming growth factor beta-1 binding protein 1) (TGF-beta1-BP-

GN 1).

GN LTPB1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE OF 1-346 FROM N.A.

RC TISSUE=Blood;

RC MEDLINE=96125117; PubMed=8537398;

RA Olafsson A., Ichijo H., Moren A., ten Dijke P., Miyazono K.,

RA Heidin C.-H.;

RT "Efficient association of an amino-terminally extended form of human

RT latent transforming growth factor-beta binding protein with the

RT extracellular matrix.";

RL J. Biol. Chem. 270:31294-31297(1995).

RL [2]

RN SEQUENCE OF 347-1595 FROM N.A.

RP TISSUE=Fibroblast, and Platelet;

RC MEDLINE=90275601; PubMed=2350783;

RA Kanzaki T., Olafsson A., Moren A., Wernstedt C., Hellman U.,

RA Miyazono K., Claesson-Welsh L., Heldin C.-H.;

RT "TGF-beta 1 binding protein: a component of the large latent complex

RT of TGF-beta 1 with multiple repeat sequences.";

RL Cell 61:1051-1061(1990).

CC -1- SUBUNIT: THE LARGE LATENT COMPLEX OF TGF-BETA1 FROM PLATELETS IS

CC COMPOSED OF THE TGF-BETA1 MOLECULE NONCOVALENTLY ASSOCIATED WITH

CC A DISULFIDE-BONDED COMPLEX OF A DIMER OF THE N-TERMINAL PROPEPTIDE

CC OF THE TGF-BETA1 PRECURSOR AND A THIRD COMPONENT DEDOTED TGF-

CC BETA1-BP. TGF-BETA1-BP DOES NOT BIND DIRECTLY TO ACTIVE TGF-BETA1.

CC -1- ALTERNATIVE PRODUCTS: 2 isoforms, a short form (AC P22064) and a

CC long form (shown here); are produced by alternative splicing.

CC -1- TISSUE SPECIFICITY: The long isoform is found in fibroblasts.

CC -1- PM: CONTAINS HYDROXYLATED ASPARAGINE RESIDUES (BY SIMILARITY).

CC -1- SIMILARITY: CONTAINS 16 EGF-LIKE DOMAINS.

CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL, LA8925; AAA96327.1; "

DR EMBL, M34057; AAA61160.1; ALT_INIT.

DR HSSP: P08709; IBB9.

DR MIM: 150390; "

DR InterPro: IPR000152; Asx_hydroxyL.

DR InterPro: IPR000561; EGF-like.

DR InterPro: IPR001881; EGF_Ca.

DR InterPro: IPR002212; Fibril-assoc.

DR Pfam: PF00008; EGF; 15.

DR Pfam: PF00683; TB; 4.

DR SMART: SMO0181; EGF_1.

DR PROSITE: PS00010; ASX_HYDROXYL; 13.

DR PROSITE: PS00022; EGF_1; 2.

DR PROSITE: PS0186; EGF_2; 11.

DR PROSITE: PS01187; EGF_CA; 15.

KW Growth factor binding; Repeat; EGF-like domain; Hydroxylation; Signal;

KW Glycoprotein; Alternative splicing.

FT SIGNAL

FT CHAIN

FT 1 23

FT 24 1595

FT POTENTIAL.

FT LATENT TRANSFORMING GROWTH FACTOR BETA

FT BINDING PROTEIN 1L.

FT EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).

FT REPEAT A.

FT EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).

FT EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).

FT EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).

FT EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN

FT 872 911

[illegible]

RESULT 7
EMBL_MYCAV STANDARD: PRT: 1065 AA.
ID EMBL_MYCAV
AC P71486;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable arabinosyltransferase B (EC 2.4.2.-).
GN EMBB.
OS Mycobacterium avium.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
OX NCBI_TaxID=1764;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2151;
RX MEDLINE=97030297; PubMed=8876238;
RA Belanger A.E., Beara G.S., Ford M.E., Mikusova K., Bellisle J.T.,
Brennan P.J., Inamine J.M.;
RT "The embAB genes of Mycobacterium avium encode an arabinosyl
transferase involved in cell wall arabinan biosynthesis that is the
target for the antimycobacterial drug ethambutol."
RL Proc. Natl. Acad. Sci. U.S.A. 93:11919-11924(1996).
CC -1- FUNCTION: Arabinosyl transferase responsible for the
polymerization of arabinose into the arabinan of arabinogalactan.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (probable).
CC -1- SIMILARITY: BELONGS TO THE EMB FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
or send an email to license@sdb-sdb.ch).
DR EMBL: U66560; AAC4548.1;
KM Transferrase, glycosyltransferase; Transmembrane; Cell wall;
KM Antibiotic resistance.
FT TRANSMEM 15
FT TRANSMEM 37
FT TRANSMEM 204
FT TRANSMEM 241
FT TRANSMEM 263
FT TRANSMEM 394
FT TRANSMEM 417
FT TRANSMEM 441
FT TRANSMEM 463
FT TRANSMEM 510
FT TRANSMEM 527
FT TRANSMEM 540
FT TRANSMEM 567
FT TRANSMEM 589
FT TRANSMEM 596
FT TRANSMEM 618
FT TRANSMEM 633
FT TRANSMEM 655
FT TRANSMEM 667
FT TRANSMEM 689
FT TRANSMEM 1065
FT TRANSMEM 114624
FT TRANSMEM 3F12D113678C62BF CRC64;
SQ SEQUENCE 1065 AA; 114624 MW; 3F12D113678C62BF CRC64;
Query Match 8.6%; Score 78.5; DB 1; Length 1065;
Best Local Similarity 23.8%; Pred. No. 13;
Matches 30; Conservative 13; Mismatches 46; Indels 37; Gaps 4;

ID M2A2_HUMAN STANDARD: PRT: 1139 AA.
AC P49641; Q13754;
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha-mannosidase IIx (EC 3.2.1.114) (Mannosyl-oligosaccharide 1,3-
DE 1,6-alpha-mannosidase) (MAN IIx) (Mannosidase alpha class 2A member
DE 2).
GN MAN2A2 OR MAN2X.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (LONG AND SHORT FORMS).
RC TISSUE=Melanoma;
RX MEDLINE=96102195; PubMed=8524845;
RA Misago M., Iiso Y.-F., Kudo S., Eco S., Mattei M.-G., Moremen K.W.,
Fukuda M.N.;
RT "Molecular cloning and expression of cDNAs encoding human alpha-
RT mannosidase II and a previously unrecognized alpha-mannosidase IIx
RT isozyme."
RL Proc. Natl. Acad. Sci. U.S.A. 92:11766-11770(1995).
CC -1- FUNCTION: CATALYZES THE FIRST COMMITTED STEP IN THE BIOSYNTHESIS
OF COMPLEX N-GLYCANS. IT CONTROLS CONVERSION OF HIGH MANNOSE TO
COMPLEX N-GLYCANS; THE FINAL HYDROLYTIC STEP IN THE N-GLYCAN
MATURATION PATHWAY.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of the terminal 1,3- and 1,6-linked
alpha-D-mannose residues in the mannosyl-oligosaccharide
Man(5)(GlcNAc)(3).
CC -1- PATHWAY: N-GLYCOSYLATION.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (by
similarity).
CC -1- ALTERNATIVE PRODUCTS: At least 2 isoforms: a long form (shown
here) and a short form; are produced by alternative splicing.
CC -1- SIMILARITY: BELONGS TO FAMILY 38 OF GLYCOSYL HYDROLASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
or send an email to license@sdb-sdb.ch).
DR EMBL: D56449; BAA09510.1;
DR EMBL: L28821; AAA92022.1;
DR GeneW: HGNC:6825; MAN2A2.
DR MIM: 600988;
DR InterPro: IPR00602; glyco_hydro_38.
DR Pfam: PF01074; glyco_hydro_38; 1.
DR Hydrolase; Glycosidase; Transmembrane; Glycoprotein; Signal-anchor;
KM Golgi stack; Alternative splicing.
FT TRANSMEM 6
FT TRANSMEM 26
FT TRANSMEM 27
FT TRANSMEM 95
FT TRANSMEM 96
FT TRANSMEM 97
FT TRANSMEM 98
FT TRANSMEM 99
FT TRANSMEM 100
FT TRANSMEM 101
FT TRANSMEM 102
FT TRANSMEM 103
FT TRANSMEM 104
FT TRANSMEM 105
FT TRANSMEM 106
FT TRANSMEM 107
FT TRANSMEM 108
FT TRANSMEM 109
FT TRANSMEM 110
FT TRANSMEM 111
FT TRANSMEM 112
FT TRANSMEM 113
FT TRANSMEM 114
FT TRANSMEM 115
FT TRANSMEM 116
FT TRANSMEM 117
FT TRANSMEM 118
FT TRANSMEM 119
FT TRANSMEM 120
FT TRANSMEM 121
FT TRANSMEM 122
FT TRANSMEM 123
FT TRANSMEM 124
FT TRANSMEM 125
FT TRANSMEM 126
FT TRANSMEM 127
FT TRANSMEM 128
FT TRANSMEM 129
FT TRANSMEM 130
FT TRANSMEM 131
FT TRANSMEM 132
FT TRANSMEM 133
FT TRANSMEM 134
FT TRANSMEM 135
FT TRANSMEM 136
FT TRANSMEM 137
FT TRANSMEM 138
FT TRANSMEM 139
FT TRANSMEM 140
FT TRANSMEM 141
FT TRANSMEM 142
FT TRANSMEM 143
FT TRANSMEM 144
FT TRANSMEM 145
FT TRANSMEM 146
FT TRANSMEM 147
FT TRANSMEM 148
FT TRANSMEM 149
FT TRANSMEM 150
FT TRANSMEM 151
FT TRANSMEM 152
FT TRANSMEM 153
FT TRANSMEM 154
FT TRANSMEM 155
FT TRANSMEM 156
FT TRANSMEM 157
FT TRANSMEM 158
FT TRANSMEM 159
FT TRANSMEM 160
FT TRANSMEM 161
FT TRANSMEM 162
FT TRANSMEM 163
FT TRANSMEM 164
FT TRANSMEM 165
FT TRANSMEM 166
FT TRANSMEM 167
FT TRANSMEM 168
FT TRANSMEM 169
FT TRANSMEM 170
FT TRANSMEM 171
FT TRANSMEM 172
FT TRANSMEM 173
FT TRANSMEM 174
FT TRANSMEM 175
FT TRANSMEM 176
FT TRANSMEM 177
FT TRANSMEM 178
FT TRANSMEM 179
FT TRANSMEM 180
FT TRANSMEM 181
FT TRANSMEM 182
FT TRANSMEM 183
FT TRANSMEM 184
FT TRANSMEM 185
FT TRANSMEM 186
FT TRANSMEM 187
FT TRANSMEM 188
FT TRANSMEM 189
FT TRANSMEM 190
FT TRANSMEM 191
FT TRANSMEM 192
FT TRANSMEM 193
FT TRANSMEM 194
FT TRANSMEM 195
FT TRANSMEM 196
FT TRANSMEM 197
FT TRANSMEM 198
FT TRANSMEM 199
FT TRANSMEM 200
FT TRANSMEM 201
FT TRANSMEM 202
FT TRANSMEM 203
FT TRANSMEM 204
FT TRANSMEM 205
FT TRANSMEM 206
FT TRANSMEM 207
FT TRANSMEM 208
FT TRANSMEM 209
FT TRANSMEM 210
FT TRANSMEM 211
FT TRANSMEM 212
FT TRANSMEM 213
FT TRANSMEM 214
FT TRANSMEM 215
FT TRANSMEM 216
FT TRANSMEM 217
FT TRANSMEM 218
FT TRANSMEM 219
FT TRANSMEM 220
FT TRANSMEM 221
FT TRANSMEM 222
FT TRANSMEM 223
FT TRANSMEM 224
FT TRANSMEM 225
FT TRANSMEM 226
FT TRANSMEM 227
FT TRANSMEM 228
FT TRANSMEM 229
FT TRANSMEM 230
FT TRANSMEM 231
FT TRANSMEM 232
FT TRANSMEM 233
FT TRANSMEM 234
FT TRANSMEM 235
FT TRANSMEM 236
FT TRANSMEM 237
FT TRANSMEM 238
FT TRANSMEM 239
FT TRANSMEM 240
FT TRANSMEM 241
FT TRANSMEM 242
FT TRANSMEM 243
FT TRANSMEM 244
FT TRANSMEM 245
FT TRANSMEM 246
FT TRANSMEM 247
FT TRANSMEM 248
FT TRANSMEM 249
FT TRANSMEM 250
FT TRANSMEM 251
FT TRANSMEM 252
FT TRANSMEM 253
FT TRANSMEM 254
FT TRANSMEM 255
FT TRANSMEM 256
FT TRANSMEM 257
FT TRANSMEM 258
FT TRANSMEM 259
FT TRANSMEM 260
FT TRANSMEM 261
FT TRANSMEM 262
FT TRANSMEM 263
FT TRANSMEM 264
FT TRANSMEM 265
FT TRANSMEM 266
FT TRANSMEM 267
FT TRANSMEM 268
FT TRANSMEM 269
FT TRANSMEM 270
FT TRANSMEM 271
FT TRANSMEM 272
FT TRANSMEM 273
FT TRANSMEM 274
FT TRANSMEM 275
FT TRANSMEM 276
FT TRANSMEM 277
FT TRANSMEM 278
FT TRANSMEM 279
FT TRANSMEM 280
FT TRANSMEM 281
FT TRANSMEM 282
FT TRANSMEM 283
FT TRANSMEM 284
FT TRANSMEM 285
FT TRANSMEM 286
FT TRANSMEM 287
FT TRANSMEM 288
FT TRANSMEM 289
FT TRANSMEM 290
FT TRANSMEM 291
FT TRANSMEM 292
FT TRANSMEM 293
FT TRANSMEM 294
FT TRANSMEM 295
FT TRANSMEM 296
FT TRANSMEM 297
FT TRANSMEM 298
FT TRANSMEM 299
FT TRANSMEM 300
FT TRANSMEM 301
FT TRANSMEM 302
FT TRANSMEM 303
FT TRANSMEM 304
FT TRANSMEM 305
FT TRANSMEM 306
FT TRANSMEM 307
FT TRANSMEM 308
FT TRANSMEM 309
FT TRANSMEM 310
FT TRANSMEM 311
FT TRANSMEM 312
FT TRANSMEM 313
FT TRANSMEM 314
FT TRANSMEM 315
FT TRANSMEM 316
FT TRANSMEM 317
FT TRANSMEM 318
FT TRANSMEM 319
FT TRANSMEM 320
FT TRANSMEM 321
FT TRANSMEM 322
FT TRANSMEM 323
FT TRANSMEM 324
FT TRANSMEM 325
FT TRANSMEM 326
FT TRANSMEM 327
FT TRANSMEM 328
FT TRANSMEM 329
FT TRANSMEM 330
FT TRANSMEM 331
FT TRANSMEM 332
FT TRANSMEM 333
FT TRANSMEM 334
FT TRANSMEM 335
FT TRANSMEM 336
FT TRANSMEM 337
FT TRANSMEM 338
FT TRANSMEM 339
FT TRANSMEM 340
FT TRANSMEM 341
FT TRANSMEM 342
FT TRANSMEM 343
FT TRANSMEM 344
FT TRANSMEM 345
FT TRANSMEM 346
FT TRANSMEM 347
FT TRANSMEM 348
FT TRANSMEM 349
FT TRANSMEM 350
FT TRANSMEM 351
FT TRANSMEM 352
FT TRANSMEM 353
FT TRANSMEM 354
FT TRANSMEM 355
FT TRANSMEM 356
FT TRANSMEM 357
FT TRANSMEM 358
FT TRANSMEM 359
FT TRANSMEM 360
FT TRANSMEM 361
FT TRANSMEM 362
FT TRANSMEM 363
FT TRANSMEM 364
FT TRANSMEM 365
FT TRANSMEM 366
FT TRANSMEM 367
FT TRANSMEM 368
FT TRANSMEM 369
FT TRANSMEM 370
FT TRANSMEM 371
FT TRANSMEM 372
FT TRANSMEM 373
FT TRANSMEM 374
FT TRANSMEM 375
FT TRANSMEM 376
FT TRANSMEM 377
FT TRANSMEM 378
FT TRANSMEM 379
FT TRANSMEM 380
FT TRANSMEM 381
FT TRANSMEM 382
FT TRANSMEM 383
FT TRANSMEM 384
FT TRANSMEM 385
FT TRANSMEM 386
FT TRANSMEM 387
FT TRANSMEM 388
FT TRANSMEM 389
FT TRANSMEM 390
FT TRANSMEM 391
FT TRANSMEM 392
FT TRANSMEM 393
FT TRANSMEM 394
FT TRANSMEM 395
FT TRANSMEM 396
FT TRANSMEM 397
FT TRANSMEM 398
FT TRANSMEM 399
FT TRANSMEM 400
FT TRANSMEM 401
FT TRANSMEM 402
FT TRANSMEM 403
FT TRANSMEM 404
FT TRANSMEM 405
FT TRANSMEM 406
FT TRANSMEM 407
FT TRANSMEM 408
FT TRANSMEM 409
FT TRANSMEM 410
FT TRANSMEM 411
FT TRANSMEM 412
FT TRANSMEM 413
FT TRANSMEM 414
FT TRANSMEM 415
FT TRANSMEM 416
FT TRANSMEM 417
FT TRANSMEM 418
FT TRANSMEM 419
FT TRANSMEM 420
FT TRANSMEM 421
FT TRANSMEM 422
FT TRANSMEM 423
FT TRANSMEM 424
FT TRANSMEM 425
FT TRANSMEM 426
FT TRANSMEM 427
FT TRANSMEM 428
FT TRANSMEM 429
FT TRANSMEM 430
FT TRANSMEM 431
FT TRANSMEM 432
FT TRANSMEM 433
FT TRANSMEM 434
FT TRANSMEM 435
FT TRANSMEM 436
FT TRANSMEM 437
FT TRANSMEM 438
FT TRANSMEM 439
FT TRANSMEM 440
FT TRANSMEM 441
FT TRANSMEM 442
FT TRANSMEM 443
FT TRANSMEM 444
FT TRANSMEM 445
FT TRANSMEM 446
FT TRANSMEM 447
FT TRANSMEM 448
FT TRANSMEM 449
FT TRANSMEM 450
FT TRANSMEM 451
FT TRANSMEM 452
FT TRANSMEM 453
FT TRANSMEM 454
FT TRANSMEM 455
FT TRANSMEM 456
FT TRANSMEM 457
FT TRANSMEM 458
FT TRANSMEM 459
FT TRANSMEM 460
FT TRANSMEM 461
FT TRANSMEM 462
FT TRANSMEM 463
FT TRANSMEM 464
FT TRANSMEM 465
FT TRANSMEM 466
FT TRANSMEM 467
FT TRANSMEM 468
FT TRANSMEM 469
FT TRANSMEM 470
FT TRANSMEM 471
FT TRANSMEM 472
FT TRANSMEM 473
FT TRANSMEM 474
FT TRANSMEM 475
FT TRANSMEM 476
FT TRANSMEM 477
FT TRANSMEM 478
FT TRANSMEM 479
FT TRANSMEM 480
FT TRANSMEM 481
FT TRANSMEM 482
FT TRANSMEM 483
FT TRANSMEM 484
FT TRANSMEM 485
FT TRANSMEM 486
FT TRANSMEM 487
FT TRANSMEM 488
FT TRANSMEM 489
FT TRANSMEM 490
FT TRANSMEM 491
FT TRANSMEM 492
FT TRANSMEM 493
FT TRANSMEM 494
FT TRANSMEM 495
FT TRANSMEM 496
FT TRANSMEM 497
FT TRANSMEM 498
FT TRANSMEM 499
FT TRANSMEM 500
FT TRANSMEM 501
FT TRANSMEM 502
FT TRANSMEM 503
FT TRANSMEM 504
FT TRANSMEM 505
FT TRANSMEM 506
FT TRANSMEM 507
FT TRANSMEM 508
FT TRANSMEM 509
FT TRANSMEM 510
FT TRANSMEM 511
FT TRANSMEM 512
FT TRANSMEM 513
FT TRANSMEM 514
FT TRANSMEM 515
FT TRANSMEM 516
FT TRANSMEM 517
FT TRANSMEM 518
FT TRANSMEM 519
FT TRANSMEM 520
FT TRANSMEM 521
FT TRANSMEM 522
FT TRANSMEM 523
FT TRANSMEM 524
FT TRANSMEM 525
FT TRANSMEM 526
FT TRANSMEM 527
FT TRANSMEM 528
FT TRANSMEM 529
FT TRANSMEM 530
FT TRANSMEM 531
FT TRANSMEM 532
FT TRANSMEM 533
FT TRANSMEM 534
FT TRANSMEM 535
FT TRANSMEM 536
FT TRANSMEM 537
FT TRANSMEM 538
FT TRANSMEM 539
FT TRANSMEM 540
FT TRANSMEM 541
FT TRANSMEM 542
FT TRANSMEM 543
FT TRANSMEM 544
FT TRANSMEM 545
FT TRANSMEM 546
FT TRANSMEM 547
FT TRANSMEM 548
FT TRANSMEM 549
FT TRANSMEM 550
FT TRANSMEM 551
FT TRANSMEM 552
FT TRANSMEM 553
FT TRANSMEM 554
FT TRANSMEM 555
FT TRANSMEM 556
FT TRANSMEM 557
FT TRANSMEM 558
FT TRANSMEM 559
FT TRANSMEM 560
FT TRANSMEM 561
FT TRANSMEM 562
FT TRANSMEM 563
FT TRANSMEM 564
FT TRANSMEM 565
FT TRANSMEM 566
FT TRANSMEM 567
FT TRANSMEM 568
FT TRANSMEM 569
FT TRANSMEM 570
FT TRANSMEM 571
FT TRANSMEM 572
FT TRANSMEM 573
FT TRANSMEM 574
FT TRANSMEM 575
FT TRANSMEM 576
FT TRANSMEM 577
FT TRANSMEM 578
FT TRANSMEM 579
FT TRANSMEM 580
FT TRANSMEM 581
FT TRANSMEM 582
FT TRANSMEM 583
FT TRANSMEM 584
FT TRANSMEM 585
FT TRANSMEM 586
FT TRANSMEM 587
FT TRANSMEM 588
FT TRANSMEM 589
FT TRANSMEM 590
FT TRANSMEM 591
FT TRANSMEM 592
FT TRANSMEM 593
FT TRANSMEM 594
FT TRANSMEM 595
FT TRANSMEM 596
FT TRANSMEM 597
FT TRANSMEM 598
FT TRANSMEM 599
FT TRANSMEM 600
FT TRANSMEM 601
FT TRANSMEM 602
FT TRANSMEM 603
FT TRANSMEM 604
FT TRANSMEM 605
FT TRANSMEM 606
FT TRANSMEM 607
FT TRANSMEM 608
FT TRANSMEM 609
FT TRANSMEM 610
FT TRANSMEM 611
FT TRANSMEM 612
FT TRANSMEM 613
FT TRANSMEM 614
FT TRANSMEM 615
FT TRANSMEM 616
FT TRANSMEM 617
FT TRANSMEM 618
FT TRANSMEM 619
FT TRANSMEM 620
FT TRANSMEM 621
FT TRANSMEM 622
FT TRANSMEM 623
FT TRANSMEM 624
FT TRANSMEM 625
FT TRANSMEM 626
FT TRANSMEM 627
FT TRANSMEM 628
FT TRANSMEM 629
FT TRANSMEM 630
FT TRANSMEM 631
FT TRANSMEM 632
FT TRANSMEM 633
FT TRANSMEM 634
FT TRANSMEM 635
FT TRANSMEM 636
FT TRANSMEM 637
FT TRANSMEM 638
FT TRANSMEM 639
FT TRANSMEM 640
FT TRANSMEM 641
FT TRANSMEM 642
FT TRANSMEM 643
FT TRANSMEM 644
FT TRANSMEM 645
FT TRANSMEM 646
FT TRANSMEM 647
FT TRANSMEM 648
FT TRANSMEM 649
FT TRANSMEM 650
FT TRANSMEM 651
FT TRANSMEM 652
FT TRANSMEM 653
FT TRANSMEM 654
FT TRANSMEM 655
FT TRANSMEM 656
FT TRANSMEM 657
FT TRANSMEM 658
FT TRANSMEM 659
FT TRANSMEM 660
FT TRANSMEM 661
FT TRANSMEM 662
FT TRANSMEM 663
FT TRANSMEM 664
FT TRANSMEM 665
FT TRANSMEM 666
FT TRANSMEM 667
FT TRANSMEM 668
FT TRANSMEM 669
FT TRANSMEM 670
FT TRANSMEM 671
FT TRANSMEM 672
FT TRANSMEM 673
FT TRANSMEM 674
FT TRANSMEM 675
FT TRANSMEM 676
FT TRANSMEM 677
FT TRANSMEM 678
FT TRANSMEM 679
FT TRANSMEM 680
FT TRANSMEM 681
FT TRANSMEM 682
FT TRANSMEM 683
FT TRANSMEM 684
FT TRANSMEM 685
FT TRANSMEM 686
FT TRANSMEM 687
FT TRANSMEM 688
FT TRANSMEM 689
FT TRANSMEM 690
FT TRANSMEM 691
FT TRANSMEM 692
FT TRANSMEM 693
FT TRANSMEM 694
FT TRANSMEM 695
FT TRANSMEM 696
FT TRANSMEM 697
FT TRANSMEM 698
FT TRANSMEM 699
FT TRANSMEM 700
FT TRANSMEM 701
FT TRANSMEM 702
FT TRANSMEM 703
FT TRANSMEM 704
FT TRANSMEM 705
FT TRANSMEM 706
FT TRANSMEM 707
FT TRANSMEM 708
FT TRANSMEM 709
FT TRANSMEM 710
FT TRANSMEM 711
FT TRANSMEM 712
FT TRANSMEM 713
FT TRANSMEM 714
FT TRANSMEM 715
FT TRANSMEM 716
FT TRANSMEM 717
FT TRANSMEM 718
FT TRANSMEM 719
FT TRANSMEM 720
FT TRANSMEM 721
FT TRANSMEM 722
FT TRANSMEM 723
FT TRANSMEM 724
FT TRANSMEM 725
FT TRANSMEM 726
FT TRANSMEM 727
FT TRANSMEM 728
FT TRANSMEM 729
FT TRANSMEM 730
FT TRANSMEM 731
FT TRANSMEM 732
FT TRANSMEM 733
FT TRANSMEM 734
FT TRANSMEM 735
FT TRANSMEM 736
FT TRANSMEM 737
FT TRANSMEM 738
FT TRANSMEM 739
FT TRANSMEM 740
FT TRANSMEM 741
FT TRANSMEM 742
FT TRANSMEM 743
FT TRANSMEM 744
FT TRANSMEM 745
FT TRANSMEM 746
FT TRANSMEM 747
FT TRANSMEM 748
FT TRANSMEM 749
FT TRANSMEM 750
FT TRANSMEM 751
FT TRANSMEM 752
FT TRANSMEM 753
FT TRANSMEM 754
FT TRANSMEM 755
FT TRANSMEM 756
FT TRANSMEM 757
FT TRANSMEM 758
FT TRANSMEM 759
FT TRANSMEM 760
FT TRANSMEM 761
FT TRANSMEM 762
FT TRANSMEM 763
FT TRANSMEM 764
FT TRANSMEM 765
FT TRANSMEM 766
FT TRANSMEM 767
FT TRANSMEM 768
FT TRANSMEM 769
FT TRANSMEM 770
FT TRANSMEM 771
FT TRANSMEM 772
FT TRANSMEM 773
FT TRANSMEM 774
FT TRANSMEM 775
FT TRANSMEM 776
FT TRANSMEM 777
FT TRANSMEM 778
FT TRANSMEM 779
FT TRANSMEM 780
FT TRANSMEM 781
FT TRANSMEM 782
FT TRANSMEM 783
FT TRANSMEM 784
FT TRANSMEM 785
FT TRANSMEM 786
FT TRANSMEM 787
FT TRANSMEM 788
FT TRANSMEM 789
FT TRANSMEM 790
FT TRANSMEM 791
FT TRANSMEM 792
FT TRANSMEM 793
FT TRANSMEM 794
FT TRANSMEM 795
FT TRANSMEM 796
FT TRANSMEM 797
FT TRANSMEM 798
FT TRANSMEM 799
FT TRANSMEM 800
FT TRANSMEM 801
FT TRANSMEM 802
FT TRANSMEM 803
FT TRANSMEM 804
FT TRANSMEM 805
FT TRANSMEM 806
FT TRANSMEM 807
FT TRANSMEM 808
FT TRANSMEM 809
FT TRANSMEM 810
FT TRANSMEM 811
FT TRANSMEM 812
FT TRANSMEM 813
FT TRANSMEM 814
FT TRANSMEM 815
FT TRANSMEM 816
FT TRANSMEM 817
FT TRANSMEM 818
FT TRANSMEM 819
FT TRANSMEM 820
FT TRANSMEM 821
FT TRANSMEM 822
FT TRANSMEM 823
FT TRANSMEM 824
FT TRANSMEM 825
FT TRANSMEM 826
FT TRANSMEM 827
FT TRANSMEM 828
FT TRANSMEM 829
FT TRANSMEM 830
FT TRANSMEM 831
FT TRANSMEM 832
FT TRANSMEM 833
FT TRANSMEM 834
FT TRANSMEM 835
FT TRANSMEM 836
FT TRANSMEM 837
FT TRANSMEM 838
FT TRANSMEM 839
FT TRANSMEM 840
FT TRANSMEM 841
FT TRANSMEM 842
FT TRANSMEM 843
FT TRANSMEM 844
FT TRANSMEM 845
FT TRANSMEM 846
FT TRANSMEM 847
FT TRANSMEM 848
FT TRANSMEM 849
FT TRANSMEM 850
FT TRANSMEM 851
FT TRANSMEM 852
FT TRANSMEM 853
FT TRANSMEM 854
FT TRANSMEM 855
FT TRANSMEM 856
FT TRANSMEM 857
FT TRANSMEM 858
FT TRANSMEM 859
FT TRANSMEM 860
FT TRANSMEM 861
FT TRANSMEM 862
FT TRANSMEM 863
FT TRANSMEM 864
FT TRANSMEM 865
FT TRANSMEM 866
FT TRANSMEM 867
FT TRANSMEM 868
FT TRANSMEM 869
FT TRANSMEM 870
FT TRANSMEM 871
FT TRANSMEM 872
FT TRANSMEM 873
FT TRANSMEM 874
FT TRANSMEM 875
FT TRANSMEM 876
FT TRANSMEM 877
FT TRANSMEM 878
FT TRANSMEM 879
FT TRANSMEM 880
FT TRANSMEM 881
FT TRANSMEM 882
FT TRANSMEM 883
FT TRANSMEM 884
FT TRANSMEM 885
FT TRANSMEM 886
FT TRANSMEM 887
FT TRANSMEM 888
FT TRANSMEM 889
FT TRANSMEM 890
FT TRANSMEM 891
FT TRANSMEM 892
FT TRANSMEM 893
FT TRANSMEM 894
FT TRANSMEM 895
FT TRANSMEM 896
FT TRANSMEM 897
FT TRANSMEM 898
FT TRANSMEM 899
FT TRANSMEM 900
FT TRANSMEM 901
FT TRANSMEM 902
FT TRANSMEM 903
FT TRANSMEM 904
FT TRANSMEM 905
FT TRANSMEM 906
FT TRANSMEM 907
FT TRANSMEM 908
FT TRANSMEM 909
FT TRANSMEM 910
FT TRANSMEM 911
FT TRANSMEM 912
FT TRANSMEM 913
FT TRANSMEM 914
FT TRANSMEM 915
FT TRANSMEM 916
FT TRANSMEM 917
FT TRANSMEM 918
FT TRANSMEM 919
FT TRANSMEM 920
FT TRANSMEM 921
FT TRANSMEM 922
FT TRANSMEM 923
FT TRANSMEM 924
FT TRANSMEM 925
FT TRANSMEM 926
FT TRANSMEM 927
FT TRANSMEM 928
FT TRANSMEM 929
FT TRANSMEM 930
FT TRANSMEM 931
FT TRANSMEM 932
FT TRANSMEM 933
FT TRANSMEM 934
FT TRANSMEM 935
FT TRANSMEM 936
FT TRANSMEM 937
FT TRANSMEM 938
FT TRANSMEM 939
FT TRANSMEM 940
FT TRANSMEM 941
FT TRANSMEM 942
FT TRANSMEM 943
FT TRANSMEM 944
FT TRANSMEM 945
FT TRANSMEM 946
FT TRANSMEM 947
FT TRANSMEM 948
FT TRANSMEM 949
FT TRANSMEM 950
FT TRANSMEM 951
FT TRANSMEM 952
FT TRANSMEM 953
FT TRANSMEM 954
FT TRANSMEM 955
FT TRANSMEM 956
FT TRANSMEM 957
FT TRANSMEM 958
FT TRANSMEM 959
FT TRANSMEM 960
FT TRANSMEM 961
FT TRANSMEM 962
FT TRANSMEM 963
FT TRANSMEM 964
FT TRANSMEM 965
FT TRANSMEM 966
FT TRANSMEM 967
FT TRANSMEM 968
FT TRANSMEM 969
FT TRANSMEM 970
FT TRANSMEM 971
FT TRANSMEM 972
FT TRANSMEM 973
FT TRANSMEM 974
FT TRANSMEM 975
FT TRANSMEM 976
FT TRANSMEM 977
FT TRANSMEM 978
FT TRANSMEM 979
FT TRANSMEM 980
FT TRANSMEM 981
FT TRANSMEM 982
FT TRANSMEM 983
FT TRANSMEM 984
FT TRANSMEM 985
FT TRANSMEM 986
FT TRANSMEM 987
FT TRANSMEM 988
FT TRANSMEM 989
FT TRANSMEM 990
FT TRANSMEM 991
FT TRANSMEM 992
FT TRANSMEM 993
FT TRANSMEM 994
FT TRANSMEM 995
FT TRANSMEM 996
FT TRANSMEM 997
FT TRANSMEM 998
FT TRANSMEM 999
FT TRANSMEM 1000
FT TRANSMEM 1001
FT TRANSMEM 1002
FT TRANSMEM 1003
FT TRANSMEM 1004
FT TRANSMEM 1005
FT TRANSMEM 1006
FT TRANSMEM 1007
FT TRANSMEM 1008
FT TRANSMEM 1009
FT TRANSMEM 1010
FT TRANSMEM 1011
FT TRANSMEM 1012
FT TRANSMEM 1013
FT TRANSMEM 1014
FT TRANSMEM 1015
FT TRANSMEM 1016
FT TRANSMEM 1017
FT TRANSMEM 1018
FT TRANSMEM 1019
FT TRANSMEM 1020
FT TRANSMEM 1021
FT TRANSMEM 1022
FT TRANSMEM 1023
FT TRANSMEM 1024
FT TRANSMEM 1025
FT TRANSMEM 1026
FT TRANSMEM 1027
FT TRANSMEM 1028
FT TRANSMEM 1029
FT TRANSMEM 1030
FT TRANSMEM 1031
FT TRANSMEM 1032
FT TRANSMEM 1033
FT TRANSMEM 1034
FT TRANSMEM 1035
FT TRANSMEM 1036
FT TRANSMEM 1037
FT TRANSMEM 1038
FT TRANSMEM 1039
FT TRANSMEM 1040
FT TRANSMEM 1041
FT TRANSMEM 1042
FT TRANSMEM 1043
FT TRANSMEM 1044
FT TRANSMEM 1045
FT TRANSMEM 1046
FT TRANSMEM 1047
FT TRANSMEM 1048
FT TRANSMEM 1049
FT TRANSMEM 1050
FT TRANSMEM 1051
FT TRANSMEM 1052
FT TRANSMEM 1053
FT TRANSMEM 1054
FT TRANSMEM 1055
FT TRANSMEM 1056
FT TRANSMEM 1057
FT TRANSMEM 1058
FT TRANSMEM 1059
FT TRANSMEM 1060
FT TRANSMEM 1061
FT TRANSMEM 1062
FT TRANSMEM 1063
FT TRANSMEM 1064
FT TRANSMEM 1065
FT TRANSMEM 1066
FT TRANSMEM 1067
FT TRANSMEM 1068
FT TRANSMEM 1069
FT TRANSMEM 1070
FT TRANSMEM 1071
FT TRANSMEM 1072
FT TRANSMEM 1073
FT TRANSMEM 1074
FT TRANSMEM 1075
FT TRANSMEM 1076
FT TRANSMEM 1077
FT TRANSMEM 1078
FT TRANSMEM 1079
FT TRANSMEM 1080
FT TRANSMEM 1081
FT TRANSMEM 1082
FT TRANSMEM 1083
FT TRANSMEM 1084
FT TRANSMEM 1085
FT TRANSMEM 1086
FT TRANSMEM 1087
FT TRANSMEM 1088
FT TRANSMEM 1089
FT TRANSMEM 1090
FT TRANSMEM 1091
FT TRANSMEM 1092
FT TRANSMEM 1093
FT TRANSMEM 1094
FT TRANSMEM 1095
FT TRANSMEM 1096
FT TRANSMEM 1097
FT TRANSMEM 1098
FT TRANSMEM 1099
FT TRANSMEM 1100
FT TRANSMEM 1101
FT TRANSMEM 1102
FT TRANSMEM 1103
FT TRANSMEM 1104
FT TRANSMEM 1105
FT TRANSMEM 1106
FT TRANSMEM 1107
FT TRANSMEM 1108
FT TRANSMEM 1109
FT TRANSMEM 1110
FT TRANSMEM 1111
FT TRANSMEM 1112
FT TRANSMEM 1113
FT TRANSMEM 1114
FT TRANSMEM 1115
FT TRANSMEM 1116
FT TRANSMEM 1117
FT TRANSMEM 1118
FT TRANSMEM 1119
FT TRANSMEM 1120
FT TRANSMEM 1121
FT TRANSMEM 1122
FT TRANSMEM 1123
FT TRANSMEM 1124
FT TRANSMEM 1125
FT TRANSMEM 1126
FT TRANSMEM 1127
FT TRANSMEM 1128
FT TRANSMEM 1129
FT TRANSMEM 1130
FT TRANSMEM 1131
FT TRANSMEM 1132
FT TRANSMEM 1133
FT TRANSMEM 1134
FT TRANSMEM 1135
FT TRANSMEM 1136
FT TRANSMEM 1137
FT TRANSMEM 1138
FT TRANSMEM 1139
FT TRANSMEM 1140
FT TRANSMEM 1141
FT TRANSMEM 1142
FT TRANSMEM 1143
FT TRANSMEM 1144
FT TRANSMEM 1145
FT TRANSMEM 1146
FT TRANSMEM 1147
FT TRANSMEM 1148
FT TRANSMEM 1149
FT TRANSMEM 1150
FT TRANSMEM 1151
FT TRANSMEM 1152
FT TRANSMEM 1153
FT TRANSMEM 1154
FT TRANSMEM 1155
FT TRANSMEM 1156
FT TRANSMEM 1157
FT TRANSMEM 1158
FT TRANSMEM 1159
FT TRANSMEM 1160
FT TRANSMEM 1161
FT TRANSMEM 1162
FT TRANSMEM 1163
FT TRANSMEM 1164
FT TRANSMEM 1165
FT TRANSMEM 1166
FT TRANSMEM 1167
FT TRANSMEM 1168
FT TRANSMEM 1169
FT TRANSMEM 1170
FT TRANSMEM 1171
FT TRANSMEM 1172
FT TRANSMEM 1173
FT TRANSMEM 1174
FT TRANSMEM 1175
FT TRANSMEM 1176
FT TRANSMEM 1177
FT TRANSMEM 1178
FT TRANSMEM 1179
FT TRANSMEM 1180
FT TRANSMEM 1181
FT TRANSMEM 1182
FT TRANSMEM 1183
FT TRANSMEM 1184
FT TRANSMEM 1185
FT TRANSMEM 1186
FT TRANSMEM 1187
FT TRANSMEM 1188
FT TRANSMEM 1189
FT TRANSMEM 1190
FT TRANSMEM 1191
FT TRANSMEM 1192
FT TRANSMEM 1193
FT TRANSMEM 1194
FT TRANSMEM 1195
FT TRANSMEM 1196
FT TRANSMEM 11

Db 863 WRGCLMTYHHWMTSGTTSWPCSTSIQTSTARVQPRRYTKLPLQANFPMVMAYI0D 922
QY 56 -----WLHYAGSSRISPTLEATLVSPFLASLRV-----ARVC--LRLLC 93
Db 923 AQRRLTLHTAQALGVSLKQGQLEV---ILDRRLM0DDNRGSLGKDKNCRTRCFRLLL 979
QY 94 PPTPKOSSTEPSNRVAMPSCPASIPQAOLMSSPRMWPYLCPLVTKITL 139
Db 980 ERTVSGEV0DSSHSTYPSLSLHLSMTYLNAP--ALALPVARMQL 1022

RESULT 9

CYSZ_SALTI
ID CYSZ_SALTI STANDARD; PRT; 253 AA.
AC 0824W3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE CysZ protein homolog.
GN CysZ OR STY265.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_Taxid=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagsels K.,
Krogh A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,
Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
Whithead S., Barrett B.G.;
RT Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar typhi CT18.";
RL Nature 413:848-852(2001).
CC -1- FUNCTION: Possibly involved in sulfate transport.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE CYSZ FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL627274; CAD07661.1; -.
CC Cysteine biosynthesis; Transmembrane; Inner membrane; Transport;
CC KW Complete proteome.
CC FT TRANSMEM 31 POTENTIAL.
CC FT TRANSMEM 68 90 POTENTIAL.
CC FT TRANSMEM 151 173 POTENTIAL.
CC FT TRANSMEM 220 242 POTENTIAL.
CC SO SEQUENCE 253 AA; 28892 MW; F1883DBD5734F906 CRC64;

Query Match 8.5%; Score 77.5; DB 1; Length 253;
Best Local Similarity 22.2%; Pred. No. 3.7;
Matches 42; Conservative 28; Mismatches 64; Indels 55; Gaps 9;

QY 8 AMIPCM--VPCWLPKRTW-----WSSSTAWVS-----WASSAL 41
Db 56 AMIPSLMSHPDWLQWLSYLLMPPLAVISVLLVGYFPSTLANWIAFPNGILAEQLRL 115
QY 42 ETSTOPATG-----ATWTKMLHYAGSSRISPTL---EATLVSPFLASLRV 84
Db 116 TGAFTPTGTGILMKDVPRIKMKRWOKLANYLRAIVLLVLFPIPGIGQITAPVLMPLFS 175

QY 85 ARVCLRLCAPPYPRKDSSTEPSNRVAMPSCPASIPQAOLMSSPRMWP-----TCLPVTKITL 139
Db 176 AMMLALGYC-DYEPDNH-----KVPFTMRALRTQVANNQFALSLTFMIVLNF 229
QY 140 RPNMAACGA 148
Db 230 MP-VAACGA 237

RESULT 10

YD71_MYCPN
ID YD71_MYCPN STANDARD; PRT; 211 AA.
AC P75410;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MPN371 (A19_orf211).
GN MPN371 OR MP465.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_Taxid=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RT Nucleic Acids Res. 24:4420-4449(1996).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000046; MAB96113.1; -.
CC Hypothetical protein; Complete proteome.
CC KW SEQUENCE 211 AA; 23592 MW; A5E24028852DDB2 CRC64;

Query Match 8.4%; Score 77; DB 1; Length 211;
Best Local Similarity 23.1%; Pred. No. 3.5;
Matches 42; Conservative 12; Mismatches 52; Indels 76; Gaps 11;

QY 15 VPCWLPKRTWMMWSSSTAWVSWASSALETSTPATGATWTKMLHYAGSSRISPTL---EA 71
Db 27 VPSWVP-----DPELEGAVPKSSALSWTCWL-----LLEPLIGALA 63
QY 72 TLTVSPFLASLR-----VARVC--LRLLCPRYKSDSTEPSW-----RVAMPs--- 112
Db 64 RLIVSSSIWPLSSSEDFEFTATCNALTLVSPDEP-----HWGWIQIOMWLKNQWPPORPG 118
QY 113 ---CPASIPQAOLMSS-----PRMW-----PTCLPVTKITLRLPMAACGARV 151
Db 119 VFHSSSCPRPKRSSSQTLPRMKWKYFDHSFAVVSPTPRATAHSTPR-----CAARVK 173
QY 152 RR 153
Db 174 RQ 175

RESULT 11

HEP_DROME
ID HEP_DROME STANDARD; PRT; 487 AA.
AC Q23977;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Dual specificity mitogen-activated protein kinase kinase hemipterous
DE (BC 2.7.1.-) (MAPKK).

GN HEP OR HEM.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-R;
 RX MEDLINE=96067643; PubMed=8521475;
 RA Glise B., Bourbon H., Nossell S.;
 RT "Hemiporous encodes a novel Drosophila MAP kinase kinase, required for epithelial cell sheet movement.";
 RL Cell 83:451-461(1995).
 CC -1- FUNCTION: REQUIRED FOR THE EPITHELIAL CELL SHEET MOVEMENT CALLED DORSAL CLOSURE (DC), WHICH ALLOWS ESTABLISHMENT OF THE DORSAL EPIDERMIS. CONTROLS THE EXPRESSION IN THE DORSAL EPITHELIAL EDGES OF ANOTHER DORSAL CLOSURE GENE, PUCKERED (PUC).
 CC -1- PFM: MAPK IS ITSELF DEPENDENT ON SER/THR PHOSPHORYLATION FOR ACTIVITY CATALYZED BY MAP KINASE KINASE KINASES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. MAP KINASE KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U05240; AAC6944.1; -;
 DR FLYBASE: FBgn0010303; hep.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase.1.
 DR ProDom: PD000001; Euk_pkinase.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM.1.
 KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase; ATP-binding; Phosphorylation; Developmental protein.
 FT DOMAIN 95 106 POLY-SER.
 FT NP_BIND 195 454 PROTEIN KINASE.
 FT BINDING 201 209 ATP (BY SIMILARITY).
 FT ACT_SITE 318 318 ATP (BY SIMILARITY).
 FT MOD_RES 346 346 BY SIMILARITY.
 FT MOD_RES 350 350 PHOSPHORYLATION (BY SIMILARITY).
 SO SEQUENCE 487 AA; 53079 MW; 09E248BD14AE5 CRC64;
 Query Match 8.4%; Score 77; DB 1; Length 487;
 Best Local Similarity 30.0%; Pred. No. 7.9;
 Matches 33; Conservative 15; Mismatches 38; Indels 24; Gaps 5;

GN AMA-1 OR F36A4.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=90066416; PubMed=2586513;
 RA Bird D.M., Riddle D.L.;
 RT "Molecular cloning and sequencing of ama-1, the gene encoding the largest subunit of Caenorhabditis elegans RNA polymerase II.";
 RL Mol. Cell. Biol. 9:4119-4130(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Miller N., Bradshaw H.;
 RT Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP REVISIONS.
 RA Waterston R.;
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + {RNA}(N).
 CC -1- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM NINE TO FOURTEEN DIFFERENT POLYPEPTIDES.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- PFM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED. THE PHOSPHORYLATION ACTIVATES POL2.
 CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE III FOR 5S AND TRNA GENES.
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M29235; AAA28126.1; -;
 DR EMBL: U53333; AAA96158.2; -;
 DR PIR: A34092; A34092.
 DR WormPep: F36A4.7; CE28300.
 DR InterPro: IPR000684; RNA_POL_II_repeat.
 DR InterPro: IPR000722; RNA_POL_A.
 DR InterPro: IPR002879; RNA_POL_A2.
 DR Pfam: PF01854; RNA_POL_A.1.
 DR PROSITE: PS00115; RNA_POL_II_REPEAT.26.
 KW Transferase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat; DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger. C2H2-TYPE (POTENTIAL).
 FT ZN_FING 66 82
 FT DOMAIN 1560 1852 CARBOXYL-TERMINAL 7-RESIDUE REPEATS.
 FT CONFLICT 215 215 V -> D (IN REF. 1).
 FT CONFLICT 911 911 R -> D (IN REF. 1).
 FT CONFLICT 959 959 I -> D (IN REF. 1).
 FT CONFLICT 974 974 Q -> L (IN REF. 1).
 FT CONFLICT 990 991 KP -> NA (IN REF. 1).
 FT CONFLICT 1156 1158 MISSING (IN REF. 1).
 FT CONFLICT 1402 1403 IT -> IS (IN REF. 1).
 SO SEQUENCE 1852 AA; 203978 MW; 211E4E563119088B CRC64;
 Query Match 8.4%; Score 77; DB 1; Length 1852;
 Best Local Similarity 30.7%; Pred. No. 29;
 Matches 31; Conservative 11; Mismatches 51; Indels 8; Gaps 3;

RESULT 15
PTPO_RAT STANDARD; PRT: 1711 AA.
AC 064612;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Osteoclastic protein tyrosine phosphatase precursor (EC 3.1.3.48)
DE (OST-PTP).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Osteosarcoma;
RX MEDLINE=95074080; PubMed=7527035;
RA Mauro L.J., Olmsted E.A., Skrobacz B.M., Mourey R.J., Davis A.R.,
RA Dixon J.E.;
RT Identification of a hormonally regulated protein tyrosine
RT phosphatase associated with bone and testicular differentiation.";
RL J. Biol. Chem. 269:30659-30667(1994).
CC -1- FUNCTION: MAY FUNCTION IN SIGNALING PATHWAYS DURING BONE
CC REMODELING. AS WELL AS SERVE A BROADER ROLE IN CELL INTERACTIONS
CC ASSOCIATED WITH DIFFERENTIATION IN BONE AND TESTIS. OPTIMAL PH FOR
CC PHOSPHATASE ACTIVITY IS 5.6.
CC ASSOCIATED WITH DIFFERENTIATION IN BONE AND TESTIS.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
CC produced by alternative splicing. A presumed alternate transcript
CC of 4.8-5.0 kilobases, which may lack PTP domains, is present in
CC proliferating osteoblasts, but not detectable at other stages.
CC -1- TISSUE SPECIFICITY: BONE AND TESTIS. IN THE LATTER, RESTRICTED TO
CC THE BASAL PORTION OF THE SEMINIFEROUS TUBULE.
CC -1- DEVELOPMENTAL STAGE: UP-REGULATED IN DIFFERENTIATING CULTURES OF
CC PRIMARY OSTEOBLASTS AND DOWN-REGULATED IN LATE STAGE MINERALIZING
CC CULTURES. IN TESTIS, EXPRESSION IS HIGHEST BETWEEN STAGES I AND
CC VII WHEN MATURING SPERMATIDS REMAIN BURIED WITHIN THE SEROLI
CC EPITHELIUM.
CC -1- INDUCTION: BY PARATHYROID HORMONE AND CYCLIC AMP ANALOGS.
CC -1- PPM: THE CYTOPLASMIC DOMAIN CONTAINS POTENTIAL PHOSPHORYLATION
CC SITES.
CC -1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
CC -1- SIMILARITY: CONTAINS 10 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L36884; AAA63911.1; .
DR HSSP: P18052; 1YFO.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR003962; FNIII_repeat.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; TYR_PP.
DR Pfam: PF00041; fn3; 7.
DR Pfam: PF00102; Y_phosphatase; 1.
DR PRINTS: PR00014; FNTYPEIII.
DR PRINTS: PR00700; PRTYPHPTASE.
DR SMART: SM00060; FN3; 6.
DR SMART: SM00194; PTPC; 1.
DR SMART: SM00012; PTPC_DSPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50055; TYR_PHOSPHATASE_PMP; 2.
KW Hydroxylase; Transmembrane; Repeat; Signal; Glycoprotein.

FT	SIGNAL	1	17	POTENTIAL.
FT	CHAIN	18	1711	OSTEOSTEITICULAR PROTEIN TYROSINE
FT	TRANSMEM	18	1074	PHOSPHATASE.
FT	DOMAIN	1075	1095	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	1096	1711	POTENTIAL.
FT	DOMAIN	125	124	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	125	215	FIBRONECTIN TYPE-III 1.
FT	DOMAIN	216	303	FIBRONECTIN TYPE-III 2.
FT	DOMAIN	304	392	FIBRONECTIN TYPE-III 3.
FT	DOMAIN	393	470	FIBRONECTIN TYPE-III 4.
FT	DOMAIN	471	562	FIBRONECTIN TYPE-III 5.
FT	DOMAIN	563	652	FIBRONECTIN TYPE-III 6.
FT	DOMAIN	653	741	FIBRONECTIN TYPE-III 7.
FT	DOMAIN	742	830	FIBRONECTIN TYPE-III 8.
FT	DOMAIN	831	921	FIBRONECTIN TYPE-III 9.
FT	DOMAIN	1150	1418	FIBRONECTIN TYPE-III 10.
FT	DOMAIN	1419	1711	PROTEIN-TYROSINE PHOSPHATASE 1.
FT	ACT_SITE	1350	1350	PROTEIN-TYROSINE PHOSPHATASE 2.
FT	CARBOHYD	42	42	BY SIMILARITY.
FT	CARBOHYD	74	74	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	89	89	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	117	117	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	174	174	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	239	239	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	259	259	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	431	431	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	570	570	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	620	620	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	649	649	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	663	663	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	737	737	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	851	851	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	882	882	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	970	970	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	982	982	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	1711 AA;	187292 MW;	0F04D2D1A47A18A0 CRC64;

Query Match 8.4%; Score 76.5; DB 1; Length 1711;
Best Local Similarity 25.2%; Pred. No. 30;
Matches 29; Conservatave 16; Mismatches 51; Indels 19; Gaps 5;

OY 26 WSSSTAMV--SWASSALESTOPATGATWTKLHYGSS-----RISPTLEATLWSP- 77
Db 296 WGPATETWTPSYSDVLV---PLRNEIMASWAGAGARGGYVLTSGPSESTLGEPE 352
OY 78 -----FLASLRVAVCRLICPPPKDSTPEPSRVAMPSCSPASTPAQMLSSPTW 127
Db 353 ECNAVFPGLPPGPHYTLQLKVLACPYDAWEGS---TWLAESALPREVPGARLM 404

Search completed: March 27, 2003, 16:11:08
Job time : 30.4737 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 16:01:02 ; Search time 47.3684 Seconds

(without alignments)
704.681 Million cell updates/sec

Title: US-10-019-219-1

Perfect score: 912
Sequence: 1 TVVRLFLALPCMMVPCWLP.....MAACGARVRRRLQTLTSLR 162

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_UHNC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	10.7	343	8	09MIAT8
2	97	10.6	332	5	096086
3	95.5	10.5	834	6	09GKQ7
4	90.5	9.9	560	3	09PAQ1
5	89.5	9.8	407	12	09ENSA4
6	87.5	9.6	135	4	09BTR89
7	87.5	9.6	790	5	020599
8	87.5	9.6	842	5	0950F5
9	87.5	9.6	880	11	09JLNT5
10	86.5	9.5	195	5	09BPS9
11	86.5	9.5	230	4	09BD64
12	86	9.4	1325	13	08QFX1
13	85.5	9.4	270	4	09H7M0
14	84.5	9.3	210	17	09Y9F6
15	84.5	9.3	551	16	09KX16
16	83.5	9.2	339	8	08WFF5

17	83.5	9.2	384	16	09PD55	09p45 xylella fas
18	83	9.1	332	16	09X077	09x077 thermotoga
19	83	9.1	377	16	092V62	092v62 rhizobium m
20	83	9.1	420	5	017057	017057 caenorhabdi
21	83	9.1	501	5	09V047	09v047 drosophila
22	83	9.1	678	5	08T1E7	08t1e7 dictyostell
23	82.5	9.0	369	3	09C275	09c275 neurospora
24	82	9.0	248	6	09N062	09n062 macaca fasc
25	82	9.0	508	4	08T8X0	08t8x0 homo sapien
26	82	9.0	5146	6	08SPM4	08spm4 bos taurus
27	81.5	8.9	254	16	09HZW3	09hzw3 pseudomonas
28	81.5	8.9	653	11	008817	008817 mus musculu
29	81	8.9	149	8	09G1C2	09g1c2 gymnotus pa
30	81	8.9	214	5	09N974	09n974 leishmania
31	81	8.9	252	4	09NX53	09nx53 homo sapien
32	81	8.9	373	5	076810	076810 anopheles g
33	81	8.9	963	4	09CC04	09cc04 homo sapien
34	80.5	8.8	152	13	091954	091954 gallus gall
35	80.5	8.8	270	4	08WVW7	08wvw7 homo sapien
36	80.5	8.8	453	5	09V8A5	09v8a5 drosophila
37	80	8.8	793	4	013876	013876 homo sapien
38	79.5	8.7	413	5	08S201	08s201 drosophila
39	79	8.7	1007	4	043379	043379 homo sapien
40	78.5	8.6	366	2	08VP07	08vp07 burkholderi
41	78.5	8.6	817	5	09N968	09n968 leishmania
42	78.5	8.6	1527	11	09SE67	09se67 rattus norv
43	78	8.6	397	4	09HB02	09hb02 homo sapien
44	78	8.6	3570	4	099552	099552 homo sapien
45	78	8.6	6315	2	09AD16	09ad16 polyantrum

ALIGNMENTS

RESULT 1

ID	Q9MIAT8	PRELIMINARY;	PRT;	343 AA.
AC	Q9MIAT8;			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	NADH dehydrogenase subunit 2.			
GN	NADH2.			
OS	Typhlonectes natans (Rubber eel).			
OG	Mitochondrion.			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Amphibia; Gymnophiona; Caeciliidae; Typhlonectes.			
OX	NCBI_TaxID=8456;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-20296752; PubMed-10835397;			
RA	Zardoya R., Meyer A.,			
RT	"Mitochondrial Evidence on the Phylogenetic Position of Caecilians			
RT	(Amphibia: Gymnophiona).";			
RL	Genetics 155:765-775(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Zardoya R.,			
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.			
CC	-1- CATALYTIC ACTIVITY: NADH + UBIDUINONE = NAD(+) + UBIDUINOL.			
DR	EMBL; AF154051; AAF78148.1; -			
DR	InterPro; IPR001750; Oxidored_q1.			
DR	Pfam; PF00361; oxidored_q1.1			
KW	Mitochondrion; NAD; Oxidoreductase; Ubiquinone.			
SO	SEQUENCE 343 AA; 37519 MW; 2BAD5B15406EB2C8 CRC64;			

Query Match 10.7%; Score 98; DB 8; Length 343;
Best Local Similarity 26.0%; Pred. No. 0.2;
Matches 32; Conservative 13; Mismatches 56; Indels 22; Gaps 4;

QY 28 SSSATWWSKASSA-----LESTQPATGATWTKW--LHAGSSRSPTLEALITVSPF 78
DB 228 SMTTAWTKAPATAATLMLVLLISLGGLPPTSGFMPKMWILHELAKQSL-PAIATTAISAL 286

OY 79 LASLRVAVCLRLCPYPKDSSTPSRWAMPSCASIPAOIMSSPRMPTCLPYTKLT 138
 DB 287 LSLFYLRLCYFTTLNQSPNSTSTPTWRKHSQSTPKLLPATM-----VTALM 334
 OY 139 LRP 141
 DB 335 LRP 337

RESULT 2

O96086 PRELIMINARY: PRT: 332 AA.
 AC 096086: 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cahepsin L-like tick cysteine proteinase B.
 OS Haemaphysalis longicornis.
 CC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
 OC Parasitiformes; Ixodida; Ixodidae; Haemaphysalis.
 OX NCBI_TaxID=44386;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HARD TICK:
 RA Mulienga A., Sugimoto C., Ingram G.M., Ohashi K., Onuma M.;
 RT "Haemaphysalis longicornis cysteine proteinase gene B.,"
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB020491; BAA34704.1; -.
 DR HSSP; P07711; 1CJL.
 DR InterPro; IPR000668; Peptidase_C1.
 DR InterPro; IPR000169; SHprot_acite.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PR00705; PAPA1N.
 DR Prodom; PD000158; Peptidase_C1; 2.
 DR PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
 KW Hydrolase; Thiol protease.
 SQ SEQUENCE 332 AA; 37001 MW; 8998FCC8D9FD2FE CRC64;

Query Match 10.6%; Score 97; DB 5; Length 332;
 Best Local Similarity 29.5%; Pred. No. 0.25;
 Matches 36; Conservative 10; Mismatches 42; Indels 34; Gaps 6;

OY 17 CULPW--RTWWSSTAVWASSALETSTOPATGATWTKMLHYAGSSRISPTLEATLT 74
 DB 158 CYRWASRTWWTARNPSA-----TTAARAASWTTPSSSTRPRASPTRGATPT 205
 OY 75 VSPFLASLRVAVCLRLCPYP--PKDSTPSRWAMPSCASIPAOIMSSPRMPTCLP 133
 DB 206 TA-----RTARATSR--PPWGPPTAGTLTSPRETRRSC-----RRLWPTVGP 246
 OY 134 VT 135
 DB 247 VS 248

RESULT 3
 O96K07 PRELIMINARY: PRT: 834 AA.
 AC 096K07: 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative heparan sulfate proteoglycan.
 GN NOYOCAN.
 OS Ovis aries (Sheep).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Szuzech S., Domowicz M., Arvanitis A., Macklin W.;
 RT "The NOYOCans: A Novel Family of Developmentally Regulated
 RT Proteoglycans."
 RL Mol. Biol. Cell 11:43a-43a(2000).
 RN [2]

RP SEQUENCE FROM N.A.
 RA Szuzech S.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF033623; AAD01973.2; -.
 DR InterPro; IPR000210; BTB_P02.
 DR Pfam; PR00651; BTB; 2.
 DR SMART; SM00225; BTB; 2.
 DR PROSITE; PS50097; BTB; 1.
 SQ SEQUENCE 834 AA; 93070 MW; 756B68FECBD8464 CRC64;

Query Match 10.5%; Score 95.5; DB 6; Length 834;
 Best Local Similarity 26.7%; Pred. No. 0.87;
 Matches 40; Conservative 13; Mismatches 42; Indels 55; Gaps 11;

OY 14 MPPCWL-----PW-----RTWWSSTAVWASSALETSTOP 47
 DB 29 VPTWMLSLGPDPMWGSPSPMRSFPRRAERMPRMSVPTTRRAM--WRSSVLRMSSTT 86
 OY 48 ATGATWTKMLHYAGS--SRISPTLEATLTVSPFLASLRVAVCLRLCPYPKDSSTPSM 106
 DB 87 QSGQRSSP-----AGTCSGSP--SAMCSGP-----CRLRLSCP-----SRP-- 124
 OY 107 RVAMPSPASLPALQMSPRMW--PTCLPY 134
 DB 125 RSLM--CPTPPPAKSAALPTSMKDPICADV 152

RESULT 4

O9P401 PRELIMINARY: PRT: 560 AA.
 AC 09P401: 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Chitinase.
 GN CTS1.
 OS Ajellomyces capsulata (Histoplasma capsulatum).
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Onygenales; Onygenaceae; Ajellomyces.
 OX NCBI_TaxID=5037;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-G217B;
 RA Thomas P.W., Cole G.T.;
 RT "Identification and functional characterization of a chitinase from
 RT Histoplasma capsulatum: protein purification and isolation of genomc
 RT and full-length cDNA."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 CC HYDROLASES).
 DR EMBL; AF159366; AAF80370.1; -.
 DR InterPro; IPR001579; Chitinase_18/2.
 DR InterPro; IPR001223; Glyco_hydro_18.
 DR Pfam; PF00704; Glyco_hydro_18; 1.
 DR Prodom; PD000471; Glyco_hydro_18; 1.
 DR PROSITE; PS01095; CHITINASE_18; 1.
 KW Glycosidase; Hydrolase.
 SQ SEQUENCE 560 AA; 61347 MW; EF34B32ED9B064E2 CRC64;

Query Match 9.9%; Score 90.5; DB 3; Length 560;
 Best Local Similarity 26.7%; Pred. No. 1.7;
 Matches 40; Conservative 13; Mismatches 56; Indels 41; Gaps 5;

OY 15 VPCWLPWR--TWWSSTAVWASSALETSTOPATGATWTKMLHY-----AGSSRISPT 68
 DB 51 IDCYVWMSSTVWVSTTTTLTLPFSASASASSPRAATSTRAADYPIYSQSESTPPPP 110
 OY 69 LEATLTVSPFLASLRVAVCLRLCPYPKDSSTPSRWAMPSCASIPAOIMSSP 125

```

Db 111 YEGSCRRP-----TPKPPSSPPIPPSPVPIRPPSPSPSHMSSP 155
QY 126 -----RMMPTCLPVTKLT 138
Db 156 DGKSIYVYVWMAIYARNYNPDLPVKLT 185

RESULT 5
Q9ENSA4 PRELIMINARY; PRT; 407 AA.
AC Q9ENSA4;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
DE Thymidine kinase.
GN TK.
OS human herpesvirus 1.
OC viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10298;
RN [1]
RP SEQUENCE FROM N.A.
RA Saijo M.;
RT "Thymidine kinase gene of acyclovir-resistant HSV-1.";
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047366; BAB11943.1; -.
DR HSSP; P03176; 2K15.
DR InterPro; IPR001889; TK_herpes.
DR Pfam; PF00693; TK_herpes; 1.
DR ProDom; PD001519; TK_herpes; 2.
KW Kinase.
SQ SEQUENCE 407 AA; 44580 MW; 8799AF63A115697 CRC64;

Query Match
Best Local Similarity 23.3%; Score 89.5; DB 12; Length 407;
Matches 44; Conservative 11; Mismatches 43; Indels 91; Gaps 11;

QY 16 PCWLPW-----RTWMW-----SSSTAWVSWASSALETSTOPATGATWTKMLHYAG 61
Db 186 PCWMSWPSRRPCPAQTSQCMGFRRITDSTAW-----PNASAPASGLTWLQMPRFA 237
QY 62 SSRSISPLEATLTVSPPLASLRVARVCLRLIC-----PPPKDSST 102
Db 238 FTG-----CLPIRCGICRAAGRGRIQDSERGRPCPR----- 270
QY 103 EPSRVAMPSCPAS-----LPAQLMSSPRMWP-----TCLPVTKLTLRPMWAACGARYKR 152
Db 271 VPSRRATRAHPDISGTRILPC--FGPPSCWPPPTATCITCLP-----GFWTSMFNASY-- 320
QY 153 RFLQLTSL 161
Db 321 ---PCTSL 326

RESULT 6
Q9BT89 PRELIMINARY; PRT; 135 AA.
AC Q9BT89;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE Similar to RIKEN CDNA 1110017H11 gene.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strauberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004290; AA04290.1; -.

```

```

SQ SEQUENCE 135 AA; 14299 MW; BA02B4131B64F00C CRC64;

Query Match
Best Local Similarity 26.5%; Score 87.5; DB 4; Length 135;
Matches 35; Conservative 8; Mismatches 40; Indels 49; Gaps 9;

QY 20 PWRTWMMSSSTAWVSWASSALETSTOPATGATWTKMLHYAGSSRISPTLEATLTVSPFL 79
Db 13 PW--WMTAEPTASCYS-----WT-----PTCSACPTRP-S 43
QY 80 ASLRVARVCLRLCPYPKDSSTEP-----SMRVAMPSCPASLPAQLMSSPRMWPCLPV 134
Db 44 SGLVTTTAC-----CTSSPTATNASSLSGPLMATAFSLGSSTP-----WEPCTP- 89
QY 135 TKLTLRPMWAAC 146
Db 90 HPLT-RPSCGTC 100

RESULT 7
Q20599 PRELIMINARY; PRT; 790 AA.
AC Q20599;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE Hypothetical 88.9 kDa protein.
GN F49E10.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Miller N.;
RT "The sequence of C. elegans cosmid F49E10.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U53341; AAC69106.1; -.
DR InterPro; IPR004294; RPE65.
DR Pfam; PF03055; RPE65; 1.
DR Hypothetical protein.
SQ SEQUENCE 790 AA; 88922 MW; B1AE384B8FECB846 CRC64;

Query Match
Best Local Similarity 9.6%; Score 87.5; DB 5; Length 790;
Matches 29; Conservative 10; Mismatches 39; Indels 39; Gaps 4;

QY 27 SSSTAWVSWASSALETSTOPATGATWTKMLHYAGSSRISPTLEATLTVSPFLASLRVAR 86
Db 627 TSTSTTTTSTATTATTPQPTTTTSK-----PYTLTQTQWTA----- 666
QY 87 VCLRLCPYPKDSSTEPVAMPSCPASLPAQLMSSPRMWPCLPVTKLTLRPMW 143
Db 667 -----PPTTVKRTTPQ-----TVPPTPKIPRMPW--LAGSGSTEDPFW 704

RESULT 8
Q950F5

```

```

ID 0950F5 PRELIMINARY; PRT; 842 AA.
AC 0950F5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 95.0 kDa protein.
GN F49E10.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RC MEDLINE=9905613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 262:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Miller N.;
RT Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
RL [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; U53341; AL06020.1; -.
DR InterPro; IPR004294; RPE65.
DR Pfam; PF03055; RPE65; 1.
KW Hypothetical protein.
SQ SEQUENCE 842 AA; 94955 MW; 3B0230FB7DB110CA CRC64;

Query Match
Best Local Similarity 24.8%; Score 87.5; DB 5; Length 842;
Matches 29; Conservative 10; Mismatches 39; Indels 39; Gaps 4;

OY 27 SSSSTAWWASSALETSQPATGATWTKWLMHYAGSSRISPLEATLVSPFLASLRVAR 86
DB 627 TSTSTTTTSTTATTTATTTQPTTTTSEK-----PYLTQTQWTWTA----- 666
OY 87 VCLRLCLPPYPKDSSTEPSRWYAWSPCAPSLPAQLMSSPRMPTCLPVTKLRLPWW 143
DB 667 -----PPTTVVKRTTPQ-----TPTTPKIRRMWP--LAGSGSTEDQPMW 704
SQ SEQUENCE

RESULT 9
O9JUL75 PRELIMINARY; PRT; 890 AA.
AC 09JUL75;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE WFS1.
GN WFS1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PANCREAS;
RC MEDLINE=21096923; PubMed=11181571;
RA Takeda K., Inoue H., Tanizawa Y., Matsuzaki Y., Oba J., Watanabe Y.,
RA Shinoda K., Oka Y.;
RT "Wfs1 (Wolfram syndrome 1) gene product: predominant subcellular
RT localization to endoplasmic reticulum in cultured cells and neuronal
RT expression in rat brain.";
RL Hum. Mol. Genet. 10:477-484(2001).

```

```

DR EMBL; AF136378; AAF61423.1; -.
DR InterPro; IPR002088; PPTA.
DR PROSITE; PS00904; PPTA; UNKNOWN_1.
SQ SEQUENCE 890 AA; 100532 MW; 657BE1019D9065B CRC64;

Query Match
Best Local Similarity 21.2%; Score 87.5; DB 11; Length 890;
Matches 28; Conservative 20; Mismatches 45; Indels 39; Gaps 5;

OY 1 TYVRLFLAWLPCMMVPCW-----LPRRTWMMSSSTAW--VSWASSALET 43
DB 633 STVKLLVMTLTAFLPCWYVRSBGMKYNSTLTWQYGLCGPRAMETNARTQIIC 692
OY 44 STQPATGATWTKWLMHYAGSSRISPLEATLVSPFLAS--LRYARVCLRLCLPPPKDSS 101
DB 693 SHLEGHRTVWTGKFKYVRTEIDNSAESALNMLPFLGDMR-----CL----- 736
OY 102 TEPSRWYAMPSC 113
DB 737 ---YGEAYPSC 744
SQ SEQUENCE

RESULT 10
O9BPS9 PRELIMINARY; PRT; 195 AA.
AC 09BPS9;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Alpha-amylase.
GN AMY-D OR CG17876.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KN-7;
RX MEDLINE=21100343; PubMed=11156987;
RA Araki H., Inomata N., Yamazaki T.;
RT "Molecular evolution of duplicated Amylase gene regions in Drosophila
RT melanogaster: Evidence of positive selection in the coding regions and
RT selective constraints in the cis-regulatory regions.";
RL Genetics 157:667-677(2001).
DR EMBL; AB043038; BAB32536.1; -.
DR HSSP; P56634; IJAE.
DR Flybase; FBgn0000078; Amy-d.
SQ SEQUENCE 195 AA; 20734 MW; 0CB0FE3660837C20 CRC64;

Query Match
Best Local Similarity 26.0%; Score 86.5; DB 5; Length 195;
Matches 34; Conservative 18; Mismatches 42; Indels 37; Gaps 9;

OY 29 SSTAW-----VSWASSALETSQPATGATWTKWLMHYAGSSRISPLEA--LTYS 76
DB 84 SPTSMRPAPETRSSSPAWSSAATPSECAP---TWT-W-----SSTWPTPEAPALAA 133
OY 77 PFLASLRVARVCLRLCLPPPKDSSSTEPSRWYAWSPCAPSLPAQLMSSPRMPTCLPVTK 136
DB 134 PPAPAPARAIPEC-----PTPRMTSTRPA-----PSATTTPTKCATAS--WSVCATLTR 180
OY 137 L--TLRPMWAA 145
DB 181 ATPCTRTWSS 191
SQ SEQUENCE

RESULT 11
O96D64 PRELIMINARY; PRT; 230 AA.
AC 096D64;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

```


DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE Hypothetical 23.8 kDa protein.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC012899; AAH12899.1; -
KW Hypothetical protein.
SQ SEQUENCE 230 AA; 23781 MW; B8F2194AC35DA358 CRC64;
Query Match 9.5%; Score 86.5; DB 4; Length 230;
Best Local Similarity 23.6%; Pred. No. 1.6;
Matches 34; Conservative 11; Mismatches 60; Indels 39; Gaps 6;
DY 9 WPCMMVPCWLPWRTMWSSTAWVSWASSALETSTPATGATWTKMLHYAGSSRISPT 68
DB 68 WVPATACSC--PQAAWYPTPTAWT--AESVLARSAGPA----- 103
DY 69 LEATLVSPFLASLRVAVOLRLCPYPKDSSTFPMRVAMPSCPASLPAQLMSSPRW 128
DB 104 ----LTPSPSTAMTTAAAGAVASCPGCAQASASWRMTVGRPAASST---APGAHWA 155
DY 129 PTCLEPYTKLRLP-----WMAAC 146
DB 156 AVCRR-SALTRAPPSCQASMLPC 178
RESULT 12
Q80FX1 PRELIMINARY; PRT; 1325 AA.
AC O80FX1;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, last sequence update)
DE RIM-binding protein 2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Hibino H., Pironkova R., Onwumere O., Vologdskaya M., Hudspeth A.J.,
RA Lesage F.;
RT "RIM-binding proteins (RBPs) couple Rab3-interacting molecules (RIMs)
to voltage-gated Ca2+ channels.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY072908; AAL67995.1; -
SQ SEQUENCE 1325 AA; 147537 MW; 1DD8C50D4CDA1B58 CRC64;
Query Match 9.4%; Score 86; DB 13; Length 1325;
Best Local Similarity 30.4%; Pred. No. 11;
Matches 35; Conservative 14; Mismatches 48; Indels 18; Gaps 4;
DY 27 SSSSTAWVSWASSALETSTPATGATWTKMLHYAGSSRIS----PTLEATLVSPFLASL 82
DB 517 STQATTIOVSKMPAL-TATGTSHGANTGCGVAKGQVAVETPTAENTLVLELMRLNL 575
DY 83 RVARVCLRLCPYPK-DSSTFPMRVAMPSCPASLPAQLMSSPRWPTCLPVR 136
DB 576 EAKEVTVRTLSAGCESYDSSV-----AATPSDLVPPSPRRAPRPAKSK 618
RESULT 13
Q9H7M0 PRELIMINARY; PRT; 270 AA.
AC Q9H7M0;
DT 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, last annotation update)
DE FLJ00050 protein (Fragment).
GN FLJ00050.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RA Ohara O., Nagase T., Kikuno R., Okumura K.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
spleen.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK024458; BAB15748.1; -
FT NON_TER 1
SQ SEQUENCE 270 AA; 30295 MW; 5736E887B6D12E5F CRC64;
Query Match 9.4%; Score 85.5; DB 4; Length 270;
Best Local Similarity 29.0%; Pred. No. 2.3;
Matches 42; Conservative 12; Mismatches 56; Indels 35; Gaps 9;
DY 19 LPRITWV--WSSSTAWVSWASSA---LETSTPATGATW--TKMLHYAGSSRISPTLE 70
DB 98 LSWSTWVEFQERLSQSWALBDSVLRNIQTSMKELTRKHMDLPSPGELERSSSVSP--- 154
DY 71 ATLVSPFLASLRVAVOLRLCPYPKDSSTFPMRVAMPSCPASL 117
DB 155 ---GIRPLECEDRLALSPPAQLGLMPSAFQDLSVLVPGGLGLSPWRAPRPSSTKGL 211
DY 118 P-----AQLMSSPR--WMPCTCLPV 134
DB 212 PONGQQAALWVAQRKIMWPCPFV 236
RESULT 14
Q9Y9F6 PRELIMINARY; PRT; 210 AA.
AC Q9Y9F6;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
DE Hypothetical protein APE2332.
OS APE2332.
GN APE2332.
OC Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=93310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatawa Y.,
RA Jin-no K., Takahashi N., Sekine M., Baba S.-I., Anai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL: AP000064; BAA81344.1; -
KW Hypothetical protein: Complete proteome.
SQ SEQUENCE 210 AA; 22143 MW; 83F83A397EF043 CRC64;
Query Match 9.3%; Score 84.5; DB 17; Length 210;
Best Local Similarity 25.0%; Pred. No. 2.2;
Matches 34; Conservative 26; Mismatches 57; Indels 19; Gaps 6;
DY 25 WSSSTAWVSWASSALETSTPATGATWTKMLHYAGSSRISPTLEATLVSPFLASLRV 84
DB 74 WSTASST--ISFKSSVDDTTTSGRTMSTPSF---SASFLAASSAPLATIMGFIPEIRL 127

OY 85 -ARVCLRLCPYPKDSST---EPSKRVAMPSCPASLPAQLMSSPRMPTCLPVTKTLR 140
DB 128 ETVACITATISAIPTFTSTETLEPPFSATMSISPTSTPPSMGT---WKS-----RTLK 178
OY 141 PMWAACGARVRRPFLQ 156
DB 179 PDFSSIRAAISSIFLQ 194

RESULT 15

O9KXI6
ID O9KXI6 PRELIMINARY; PRT; 551 AA.
AC O9KXI6:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative integral membrane protein.
CN SC02334 OR SC053.25.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID-1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Ruter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)."
RL Nature 417:141-147(2002).
DR EMBL; AL357591; CAB93454.1;
SQ SEQUENCE 551 AA; 59486 MW; 6DF4A5BE234D8E9E CRC64;

Query Match 9.3%; Score 84.5; DB 16; Length 551;

Best Local Similarity 24.7%; Pred No. 5.9; Mismatches 46; Indels 73; Gaps 9;

OY 7 IAMLPCMVPCWLPKRTWMSSTST-----AMVSWASSALETSTOPATGATWTKWL 57
DB 35 LATLPT-----IPLYVVMATFLATGGDLAQAQEAADPFS---RHGSAVGLFWYCGM 84
OY 58 HTAGSSRISPTLEA-----TLTV-----SPTLASLVARVCLRLCPYPKDSSTEPSNR 107
DB 85 HTANYSVISPYLMAVGVRTVTVSGLAASWLAVALVYRCGYR-----R 128
OY 108 VAMPSCPASLPAQLMSSPRMW-----PTCLPVTKTLRPMWAACG 147
DB 129 PWPALLASLAL-----WCDVASGRATFALGVALALAACVPLVR--ERRIMLAAG 176

Search completed: March 27, 2003, 16:12:05
Job time : 50.3684 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 15:54:52 ; Search time 5,42105 Seconds
(without alignments)
221.222 Million cell updates/sec

Title: US-10-019-219-2
Perfect score: 63
Sequence: 1 SPRMPPTCL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSeq/101002:*
2: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
3: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
5: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
6: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
7: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
8: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
9: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
10: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
11: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
12: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
13: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
14: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
15: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
16: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
17: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
18: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
19: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
20: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
21: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
22: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
23: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
24: /SIDSeq/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	63	100.0	9	22	AA1980	Peptide fragment o
2	63	100.0	162	22	AA1981	Protein encoded by
3	63	100.0	166	22	AA1982	Human cardiovacu
4	46	73.0	72	22	AA1983	Vpr binding protei
5	46	73.0	105	19	AA1984	Human foetal prote
6	46	73.0	105	19	AA1985	Human foetal prote
7	46	73.0	102	22	AA1986	Human foetal prote
8	45	71.4	107	22	AA1987	Human foetal prote
9	44	69.8	136	21	AA1988	Human foetal prote
10	43	68.3	49	20	AA1989	HCV NS5B carboxy-t

11	43	68.3	237	22	AA1990	Human protein sequ
12	43	68.3	238	22	AA1991	Human purified sec
13	43	68.3	817	19	AA1992	Human discs large
14	43	68.3	817	20	AA1993	Amino acid sequenc
15	43	68.3	849	19	AA1994	Human discs large
16	43	68.3	3010	20	AA1995	Infectious hepatitis
17	43	68.3	3010	22	AA1996	Amino acid sequenc
18	43	68.3	3010	22	AA1997	Protein encoded by
19	42	66.7	34	22	AA1998	Peptide #11306 enc
20	42	66.7	34	22	AA1999	Protein #8737 enco
21	42	66.7	34	22	AA2000	Human brain expres
22	42	66.7	34	22	AA2001	Human bone marrow
23	42	66.7	34	22	AA2002	Peptide #7893 enco
24	42	66.7	34	22	AA2003	Peptide #11745 enc
25	42	66.7	34	22	AA2004	Human peptide enco
26	42	66.7	58	22	AA2005	Proteinolabacterium
27	42	66.7	112	22	AA2006	Human EST encoded
28	42	66.7	153	23	AA2007	Human ovarian anti
29	41	65.1	50	22	AA2008	Proteinolabacterium
30	41	65.1	58	22	AA2009	Proteinolabacterium
31	41	65.1	68	21	AA2010	Human secreted pro
32	41	65.1	94	22	AA2011	Proteinolabacterium
33	41	65.1	96	22	AA2012	Novel human diago
34	41	65.1	124	22	AA2013	Human nervous syst
35	41	65.1	124	22	AA2014	Human ORFX protein
36	41	65.1	146	22	AA2015	Human immune/hema
37	40	63.5	96	22	AA2016	Proteinolabacterium
38	40	63.5	106	22	AA2017	Human immune/hema
39	40	63.5	155	20	AA2018	Human yb37.1 secre
40	40	63.5	155	22	AA2019	Human secreted pro
41	40	63.5	155	23	AA2020	Human polypeptide
42	40	63.5	158	22	AA2021	Human immune/hema
43	40	63.5	293	21	AA2022	Eucalyptus grandis
44	40	63.5	447	22	AA2023	Novel human diago
45	40	63.5	518	23	AA2024	Arabidopsis transc

ALIGNMENTS

RESULT 1	
AA1980	Peptide fragment of a human intestinal carboxylesterase (ICE).
AA1981	Protein encoded by
AA1982	Human cardiovacu
AA1983	Vpr binding protei
AA1984	Human foetal prote
AA1985	Human foetal prote
AA1986	Human foetal prote
AA1987	Human foetal prote
AA1988	Human foetal prote
AA1989	HCV NS5B carboxy-t
AA1990	Human protein sequ
AA1991	Human purified sec
AA1992	Human discs large
AA1993	Amino acid sequenc
AA1994	Human discs large
AA1995	Infectious hepatitis
AA1996	Amino acid sequenc
AA1997	Protein encoded by
AA1998	Peptide #11306 enc
AA1999	Protein #8737 enco
AA2000	Human brain expres
AA2001	Human bone marrow
AA2002	Peptide #7893 enco
AA2003	Peptide #11745 enc
AA2004	Human peptide enco
AA2005	Proteinolabacterium
AA2006	Human EST encoded
AA2007	Human ovarian anti
AA2008	Proteinolabacterium
AA2009	Proteinolabacterium
AA2010	Human secreted pro
AA2011	Proteinolabacterium
AA2012	Novel human diago
AA2013	Human nervous syst
AA2014	Human ORFX protein
AA2015	Human immune/hema
AA2016	Proteinolabacterium
AA2017	Human immune/hema
AA2018	Human yb37.1 secre
AA2019	Human secreted pro
AA2020	Human polypeptide
AA2021	Human immune/hema
AA2022	Eucalyptus grandis
AA2023	Novel human diago
AA2024	Arabidopsis transc

PS Claim 2; Page 33; 53p; French.

XX The present sequence is derived from a human intestinal carboxylesterase

CC (ICE) polypeptide. ICE induces specific cytotoxic T lymphocytes

CC (CTL) and secretion by these CTL of cytotoxic factors, e.g.

CC Interleukin-2, interferon gamma and tumour necrosis factor. ICE

CC polypeptides and polynucleotides are used for treating cancer, by in

CC vivo or in vitro immunisation, particularly solid cancers and most

CC especially hepatocarcinoma or adenocarcinoma of colon and kidney. They

CC are also used to stimulate the immune system, and to increase, in

CC culture, the production of associated-associated CTL, for reinjection,

CC and/or to induce secretion of cytotoxic factors from CTL. Dendritic

CC cells loaded with ICE are used to induce such CTL in cultures.

XX

SO Sequence 9 AA;

Query Match 100.0%; Score 63; DB 22; Length 9;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SPRMWPCTL 9

DB 1 SPRMWPCTL 9

|||||

RESULT 2

AAB31701

ID AAB31701 standard; peptide; 162 AA.

XX AAB31701;

AC

XX

XX 30-APR-2001 (first entry)

DT

XX

XX Peptide fragment of a human intestinal carboxylesterase (ICE).

DE

XX

XX Intestinal carboxylesterase; ICE; tumour; cytotoxic T lymphocyte; CTL;

KW cytotoxic factor; interleukin-2; interferon gamma; adenocarcinoma;

KW tumour necrosis factor; cancer; hepatocarcinoma; colon; kidney.

XX

OS Homo sapiens.

XX

PN WO200100784-A2.

PD

XX

XX 04-JAN-2001.

PF

XX

XX 27-JUN-2000; 2000WO-FR01791.

PR

XX

XX 28-JUN-1999; 99FR-0008224.

PA (INSR) INST ROUSSY GUSTAVE.

XX

XX Ronsin C, Scott V, Triebel F;

PI

XX

XX WPI; 2001-112443/12.

DR

XX

XX New peptides and its encoding nucleic acid derived from intestinal

PT carboxylesterase, useful as immunostimulants for treating cancer -

XX

PS Claim 1; Page 3; 53p; French.

XX

CC The present sequence is derived from a human intestinal carboxylesterase

CC (ICE) polypeptide. ICE induces specific cytotoxic T lymphocytes

CC (CTL) and secretion by these CTL of cytotoxic factors, e.g.

CC Interleukin-2, interferon gamma and tumour necrosis factor. ICE

CC polypeptides and polynucleotides are used for treating cancer, by in

CC vivo or in vitro immunisation, particularly solid cancers and most

CC especially hepatocarcinoma or adenocarcinoma of colon and kidney. They

CC are also used to stimulate the immune system, and to increase, in

CC culture, the production of associated-associated CTL, for reinjection,

CC and/or to induce secretion of cytotoxic factors from CTL. Dendritic

CC cells loaded with ICE are used to induce such CTL in cultures.

XX

SO Sequence 162 AA;

Query Match 100.0%; Score 63; DB 22; Length 162;

Best Local Similarity 100.0%; Pred. No. 0.026;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SPRMWPCTL 9

DB 124 SPRMWPCTL 132

|||||

RESULT 3

AAB31703

ID AAB31703 standard; Protein; 166 AA.

XX

XX AAB31703;

AC

XX

XX 30-APR-2001 (first entry)

DT

XX

XX Protein encoded by an intestinal carboxylesterase (ICE) cDNA.

DE

XX

XX Intestinal carboxylesterase; ICE; tumour; cytotoxic T lymphocyte; CTL;

KW cytotoxic factor; interleukin-2; interferon gamma; adenocarcinoma;

KW tumour necrosis factor; cancer; hepatocarcinoma; colon; kidney.

XX

OS Homo sapiens.

XX

PN WO200100784-A2.

PD

XX

XX 04-JAN-2001.

PF

XX

XX 27-JUN-2000; 2000WO-FR01791.

PR

XX

XX 28-JUN-1999; 99FR-0008224.

PA (INSR) INST ROUSSY GUSTAVE.

XX

XX Ronsin C, Scott V, Triebel F;

PI

XX

XX WPI; 2001-112443/12.

DR

XX

XX N-PSDB; AAF25258.

PT

XX

XX New peptides and its encoding nucleic acid derived from intestinal

PT carboxylesterase, useful as immunostimulants for treating cancer -

XX

PS Disclosure; Fig 8A; 53p; French.

XX

CC The present sequence is encoded by the coding region of human intestinal

CC carboxylesterase (ICE) gene. ICE induces specific cytotoxic T

CC lymphocytes (CTL) and secretion by these CTL of cytotoxic factors,

CC e.g. interleukin-2, interferon gamma and tumour necrosis factor. ICE

CC polypeptides and polynucleotides are used for treating cancer, by in

CC vivo or in vitro immunisation, particularly solid cancers and most

CC especially hepatocarcinoma or adenocarcinoma of colon and kidney. They

CC are also used to stimulate the immune system, and to increase, in

CC culture, the production of associated-associated CTL, for reinjection,

CC and/or to induce secretion of cytotoxic factors from CTL. Dendritic

CC cells loaded with ICE are used to induce such CTL in cultures.

XX

SO Sequence 166 AA;

Query Match 100.0%; Score 63; DB 22; Length 166;

Best Local Similarity 100.0%; Pred. No. 0.027;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SPRMWPCTL 9

DB 128 SPRMWPCTL 136

|||||

RESULT 4

AAU21874

ID AAU21874 standard; Protein; 72 AA.

XX

AC AAU21874;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human cardiovascular system antigen polypeptide SEQ ID No 648.
XX
KW Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;
KW Chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; neotropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnereary; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; tissue regeneration;
KW anti-fertility.
XX
OS Homo sapiens.
XX
PN WO20015321-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01340.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.

PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI: 2001-451930/48.
 DR N-PSDB: AAS35148.
 XX
 PT New cardiovascular system related polynucleotides and polypeptides,
 PT useful for diagnosing, treating and/or preventing disorders of the
 PT cardiovascular system -
 PS
 PS Claim 11: SEQ ID NO 648; 674pp; English.
 XX
 CC Sequences AAV21852-AAU22466 represent the cardiovascular system antigen
 CC polypeptides of the invention. Cardiovascular system antigens and their
 CC associated polynucleotides are useful in the diagnosis, treatment and
 CC prevention of various types of disorders in e.g. humans, mice, rabbits,
 CC goats, horses, cats, dogs, chickens or sheep. A pathological condition
 CC can be determined by detecting the presence or absence of a mutation in a
 CC cardiovascular system antigen polynucleotide. The treatable disorders
 CC include autoimmune diseases such as Rheumatoid arthritis,
 CC hyperproliferative disorders such as neoplasms of the breast or liver,
 CC cardiovascular disorders such as cardiac arrest, cerebrovascular
 CC disorders such as cerebral ischaemia, nervous system disorders such as
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,
 CC ocular disorders such as corneal infection, endocrine disorders such as
 CC premature labour and infertility, gastrointestinal disorders such as
 CC Crohn's disease, renal disorders such as glomerulonephritis and
 CC respiratory disorders such as asthma and pleurisy. The polypeptides can
 CC also be used to aid wound healing, to prevent skin aging due to sunburn,
 CC to maintain organs before transplantation, to regenerate tissues and in
 CC chemotaxis.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

Query Match 73.0%; Score 46; DB 22; Length 72;
 Best Local Similarity 85.7%; Pred. NO. 4.6;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RMMPTCL 9
 |||||
 Db 49 RMMPTCL 55

RESULT 5
 AAW68196
 ID AAW68196 standard; peptide: 105 AA.
 XX
 AC AAW68196;
 XX
 DT 29-OCT-1998 (first entry)
 XX
 DE Vpr binding protein B29-1 amino acid sequence.
 XX
 KW Lentivirus; Vpr protein; HIV infection; cell stasis; cell death;

KW cancer; autoimmune disease; B29-1.
 XX
 OS Homo sapiens.
 XX
 PN WO9835032-A2.
 XX
 PD 13-AUG-1998.
 XX
 PF 11-FEB-1998; 98WO-US03390.
 XX
 PR 24-OCT-1997; 97US-0959279.
 PR 11-FEB-1997; 97US-0798597.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Chen ISY, Dieckmann T, Felgou J, Jowett JBM, Pogh B;
 PI Stewart SA, Withers-Ward E;
 XX
 DR WPI: 1998-447229/38.
 XX
 PT Arresting cell growth using lentivirus Vpr virion protein - used for
 PT treatment of cancer and screening for agents that reduce Vpr
 PT binding, e.g. anti-HIV agents
 PS
 PS Disclosure; Fig 7; 71pp; English.
 XX
 CC This represents the amino acid sequence of the B29-1 protein. This
 CC protein can bind to the Vpr protein encoded by the HIV genome. This is
 CC used as a cellular target in the method of the invention of identifying
 CC antitumour therapeutic candidates. The invention provides a method for
 CC arresting the growth of a cell by treatment with a Vpr lentivirus
 CC protein or its analogue. Agents that reduce binding of Vpr to a cellular
 CC target are useful for treating HIV (human immune deficiency virus)
 CC infection or more generally for restoring growth. The antitumour agent
 CC identified is useful for treating any type of cancer, since it induces
 CC cell stasis (blocks development at the G2 stage) and death. The agents
 CC can also be used for treating autoimmune diseases.
 XX
 SQ Sequence 105 AA;

Query Match 73.0%; Score 46; DB 19; Length 105;
 Best Local Similarity 75.0%; Pred. NO. 6.6;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPRMPTC 8
 |||||
 Db 60 SSRMPTC 67

RESULT 6
 AAW61569
 ID AAW61569 standard; peptide: 105 AA.
 XX
 AC AAW61569;
 XX
 DT 27-OCT-1998 (first entry)
 XX
 DE Vpr protein binding B29-1 amino acid sequence.
 XX
 KW Lentivirus; Vpr protein; HIV infection; cell stasis;
 KW cell death.
 XX
 OS Homo sapiens.
 OS
 PN WO9835234-A1.
 XX
 PD 13-AUG-1998.
 XX
 PF 11-FEB-1998; 98WO-US03008.
 XX
 PR 11-FEB-1997; 97US-0797907.
 XX
 PA (REGC) UNIV CALIFORNIA.

XX Chen ISY, Jowett JBM, Withers-ward E;
XX WPI; 1998-447375/38.
XX
PT Identification of compounds binding the HIV-1 Vpr protein - that
PT block Vpr-mediated cell stasis, useful for treating HIV-1 or other
PT lentiviral infections
XX
XX Disclosure; Fig 8; 63pp; English.
XX
CC This represents the amino acid sequence of the B29-1 protein. This
CC protein can bind to the Vpr protein encoded by the HIV genome. The
CC invention provides a method of identifying an agent for use in treating
CC lentiviral infections. The method comprises contacting a cellular target
CC of the Vpr protein with the agent to be tested, and assessing the ability
CC of the agent to block interaction of the Vpr protein with the cellular
CC target, where an agent which blocks this interaction is an anti-
CC lentiviral agent. Alternatively, the agent contacts a cell expressing the
CC Vpr protein under conditions where the Vpr protein induces cell stasis in
CC the absence of the agent. Identification of the agent is then observed
CC by blockage of Vpr-induced cell stasis. The method allows the
CC identification of compounds that block Vpr-mediated cell stasis and
CC ultimately cell death. The compounds can thus be used in the treatment
CC of HIV and other lentiviral infections.
XX
SQ Sequence 105 AA:

Query Match 73.0%; Score 46; DB 19; Length 105;
Best Local Similarity 75.0%; Pred. No. 6.6;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPRMWPYC 8
| | | | |
Db 60 SSRMWPAC 67

RESULT 7
AAM06407
ID AAM06407 standard; Protein; 102 AA.
XX
AC AAM06407;
XX
DT 05-OCT-2001 (first entry)
XX
DE Human foetal protein, SEQ ID NO: 138.
XX
KW Human; foetal protein; cytostatic; immunosuppressive; immunostimulant;
KW neurotropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;
KW gene therapy; antisense therapy; cancer; immune disorder;
KW growth disorder; osteoporosis; thrombolytic disorder;
KW nervous system disorder; inflammation.
XX
OS Homo sapiens.
XX
PN WO200155339-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US02723.
XX
PR 25-JAN-2000; 2000US-0491404.
PR 15-SEP-2000; 2000US-0663870.
PR 06-NOV-2000; 2000US-0707351.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Yeung G, Ford JE, Boyle BJ, Arterburn MC, Drmanac RA, Tang YT;
PI Liu C, Asundi V, Zhou P, Werhman T;
XX
DR WPI; 2001-465571/50.
DR N-PSDB; AAH94082.
XX

PT Novel fetal proteins useful for the treatment and diagnosis of diseases
PT associated with dysfunction of the protein e.g. cancers, immune
PT disorders, growth disorders, thrombolytic disorders, nervous system
PT disorders and inflammation -
XX
PS Claim 10; Page 231; 715pp; English.
XX
CC The invention relates to novel foetal polypeptides encoded by
CC polynucleotides comprising one of 477 sequences fully defined in the
CC specification. The foetal polynucleotides and polypeptides are
CC useful in the treatment and diagnosis of diseases such as cancers,
CC immune disorders, growth disorders (e.g. osteoporosis), thrombolytic
CC disorders, nervous system disorders and inflammation. The present
CC sequence is a polypeptide encoded by a cDNA assembled using
CC an expressed sequence tag (EST) found to be expressed in human
CC foetal tissue cDNA libraries.
XX
SQ Sequence 102 AA:

Query Match 71.4%; Score 45; DB 22; Length 102;
Best Local Similarity 75.0%; Pred. No. 9.1;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPRMWPYC 8
| | | | |
Db 16 SSRMWPYC 23

RESULT 8
AAM06855
ID AAM06855 standard; Protein; 107 AA.
XX
AC AAM06855;
XX
DT 05-OCT-2001 (first entry)
XX
DE Human foetal protein, SEQ ID NO: 1063.
XX
KW Human; foetal protein; cytostatic; immunosuppressive; immunostimulant;
KW neurotropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;
KW gene therapy; antisense therapy; cancer; immune disorder;
KW growth disorder; osteoporosis; thrombolytic disorder;
KW nervous system disorder; inflammation.
XX
OS Homo sapiens.
XX
PN WO200155339-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US02723.
XX
PR 25-JAN-2000; 2000US-0491404.
PR 15-SEP-2000; 2000US-0663870.
PR 06-NOV-2000; 2000US-0707351.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Yeung G, Ford JE, Boyle BJ, Arterburn MC, Drmanac RA, Tang YT;
PI Liu C, Asundi V, Zhou P, Werhman T;
XX
DR WPI; 2001-465571/50.
DR N-PSDB; AAH94530.
XX
PT Novel fetal proteins useful for the treatment and diagnosis of diseases
PT associated with dysfunction of the protein e.g. cancers, immune
PT disorders, growth disorders, thrombolytic disorders, nervous system
PT disorders and inflammation -
XX
PS Example 4; Page 600-601; 715pp; English.
XX
CC The invention relates to novel foetal polypeptides encoded by
CC polynucleotides comprising one of 477 sequences fully defined in the

CC specification. The foetal polynucleotides and polypeptides are
CC useful in the treatment and diagnosis of diseases such as cancers,
CC immune disorders, growth disorders (e.g. osteoporosis), thrombolytic
CC disorders, nervous system disorders and inflammation. The present
CC sequence is a polypeptide encoded by a cDNA assembled using
CC an expressed sequence tag (EST) found to be expressed in human
CC foetal tissue cDNA libraries.
XX
SO Sequence 107 AA;

Query Match 71.4%; Score 45; DB 22; Length 107;
Best Local Similarity 75.0%; Pred. No. 9.5;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPRWMPYC 8
DB 16 SRRWMPYC 23

RESULT 9
AAG22709
ID AAG22709 standard; Protein; 136 AA.
XX
AC AAG22709;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 25743.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
FE 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 69.8%; Score 44; DB 21; Length 136;
Best Local Similarity 85.7%; Pred. NO. 17;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRWPTC 8
|||||
Db 24 PRWPTC 30

RESULT 10
AAV25308
ID AAV25308 standard; Protein; 49 AA.

AC AAV25308;
XX 03-SEP-1999 (first entry)
DT
XX HCV NS5B carboxy-terminus protein fragment from genotype HCV J491.

DE
XX NS5B; antiviral compound; truncated protein; mutant; immunoprotective;
KW vaccine; immunological response; protection; disease; Flaviviridae;
KW virus; x-ray crystallography; hydrophobic tail.
XX
OS Hepatitis C virus.

PN MO9929843-A1.

PD 17-JUN-1999.

PF 09-DEC-1998; 98WO-US26070.

PR 11-DEC-1997; 97US-0069208.

PA (SMIR) SMITHKLINE BEECHAM CORP.

PI Del Vecchio A;

DR WPI; 1999-404939/34.

PT Nucleic acid encoding Hepatitis C Virus NS5B truncated protein
XX Disclosure; Page 16; 63pp; English.

CC This invention describes a novel nucleic acid encoding a Hepatitis C
CC Virus (HCV) truncation mutant of NS5B. The product of the invention has
CC immunoprotective activity and can be used as a vaccine. the NS5B mutant
CC or a nucleic acid vector directing expression of the mutant, are
CC used to induce an immunological response in a mammal. Antibodies against
CC the NS5B mutant are used to protect mammals against diseases caused by
CC viruses of the Flaviviridae. Additionally, soluble protein produced by
CC the method would allow for determination of the structure of the protein
CC via x-ray crystallography or other known methods. The HCV NS5B
CC truncation mutant has a deletion of a hydrophobic tail which release the
CC protein into the soluble portion of the cell, allowing for a greater
CC recovery of soluble protein for screening for inhibitors of NS5B
CC enzymatic activity. AAV25308-Y25322 represent the carboxy-terminus of the
CC NS5B protein from various HCV genotypes which are used to describe the
CC method of the invention.
CC
XX
SQ Sequence 49 AA;

Query Match 68.3%; Score 43; DB 20; Length 49;
Best Local Similarity 75.0%; Pred. No. 9.2;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRWPTC 9
|||||
Db 27 PRWPTC 34

RESULT 11
AAB95763
ID AAB95763 standard; Protein; 237 AA.
XX
AC AAB95763;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:18691.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
KW Homo sapiens.
XX
PN EPI074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakematsu A, Nagai K, Otsuki T;
XX
DR WPI: 2001-318749/34.
XX
PT primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 18691; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95693 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 237 AA;
Query Match 68.3%; Score 43; DB 22; Length 237;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 WMPYC 8
IIIIII
Db 99 WMPYC 103

RESULT 12
AAU69468
ID AAU69468 standard; Protein; 238 AA.
XX
AC AAU69468;
XX
DT 29-JAN-2002 (first entry)
XX
DE Human purified secretory polypeptide #37.
XX
KW Human; purified secretory polypeptide; cell proliferative disorder;
KW cancer; immune system disorder; neurological disorder; mental disorder;
KW motor neuron disorder; demyelinating disorder; neuromuscular disorder;
KW central nervous system disorder; enzyme linked immunosorbent assay;
KW ELISA; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200162918-A2.
XX
PD 30-AUG-2001.
XX
PF 01-FEB-2001; 2001WO-US03465.
XX
PR 24-FEB-2000; 2000US-185215P.
PR 24-FEB-2000; 2000US-185216P.
PR 16-MAY-2000; 2000US-205232P.
PR 17-MAY-2000; 2000US-205286P.
PR 17-MAY-2000; 2000US-205287P.
PR 17-MAY-2000; 2000US-205323P.
PR 17-MAY-2000; 2000US-205324P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;
PI Chen A, D'sa SA, Amshay S, Dahl CR, Dam TC, Daniels SE;
PI Dourou GE, Flores V, Fong WT, Greenwall LB, Hillman JL, Jones AL;
PI Liu TF, Roseberry AM, Rosen BH, Russo PD, Stockdrehner TK, Daffo A;
PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;
PI Cohen HJ, Hodgson DM, Lincoln SE;
XX
DR WPI: 2001-648217/74.
XX
PT Nucleic acids encoding secretory polypeptides, useful in genetic
PT diagnosis and therapy -
XX
PS Disclosure: Page 211-212; 237pp; English.
XX
CC Sequences AAU69432-AAU69511 represent purified secretory polypeptides of
CC the invention. The polypeptides and their associated polynucleotides can
CC be used in the treatment, prevention and diagnosis of diseases associated
CC with inappropriate secretory protein expression. These diseases include
CC cell proliferative disorders such as atherosclerosis and psoriasis,
CC cancers such as leukaemia and melanoma, immune system disorders such as
CC asthma and diabetes mellitus, neurological disorders such as epilepsy and
CC Parkinson's disease, mental disorders such as schizophrenia and seasonal
CC affective disorder (SAD), motor neuron disorders such as amyotrophic
CC lateral sclerosis, demyelinating disorders such as multiple sclerosis,
CC central nervous system disorders such as mental retardation and
CC neurofibromatosis and neuromuscular disorders such as cerebral palsy and
CC muscular dystrophy. Target polynucleotides in a sample can be detected by
CC hybridising the sample with a probe sequence complementary to the target
CC polynucleotide, under conditions in which a hybridisation complex is
CC formed, and detecting the presence or absence of the complex. The
CC polypeptides may also be used as antigens in the production of antibodies
CC against secretory proteins and in assays to identify modulators of
CC protein expression and activity. The antibodies may also be used as
CC diagnostic agents for detecting the presence of the sequences of the
CC invention in samples e.g. by enzyme linked immunosorbent assay (ELISA).
XX
SQ Sequence 238 AA;

Query Match 68.3%; Score 43; DB 22; Length 238;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 PRWPTC 8
 |||||
 DB 100 PRWPTC 104

RESULT 13

AAW48101
 ID AAW48101 standard; Protein; 817 AA.

AC AAW48101;

DT 15-JUN-1998 (first entry)

DE Human discs large 1 gene-1 cancer related molecule.

KW Human; NE-dlg; discs large 1 gene; cancer related molecule; nerve;
 internal secretion tissue.

OS Homo sapiens.

PN JP10066581-A.

PD 10-MAR-1998.

PF 23-AUG-1996; 96JP-0241370.

PR 23-AUG-1996; 96JP-0241370.

PA (SUME) SUMITOMO ELECTRIC IND CO.

DR WPI; 1998-224339/20.

DR N-PSDB; AAV20419.

PT Human discs large 1 gene family - useful in, e.g. therapeutic
 composition(s) for treating cancer

PS Claim 7; Page 14-16; 31pp; Japanese.

CC The present sequence represents human dlg-1 molecule. The present
 CC invention describes human discs large 1 gene (dlg) family expressible in
 CC nerve tissue. Also described are: (1) a polynucleotide (PN) encoding dlg
 CC and comprising a 3100 (I) or 3213 bp (II) sequence; (2) RNA corresponding
 CC to PN having sequences (I) and (II); (3) RNA having a base sequence
 CC translated by dlg; (4) an antisense PN having a at least 15 bp sequence,
 CC and which is a part of PN of (1); (5) a derivative of the antisense PN;
 CC (6) an antibody specific for dlg, and (7) an antibody specific for a
 CC polypeptide having a 817 or 849 aa sequence. The polypeptide, RNA's and
 CC antibodies can be used for detection of dlg. The antisense PN can be
 CC used as a therapeutic composition for treating cancer.

SQ Sequence 817 AA;

Query Match 68.3%; Score 43; DB 19; Length 817;
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRWPTC 8
 |||||

DB 103 PSMWPEC 109

RESULT 14

AA30137
 ID AAV30137 standard; Protein; 817 AA.

AC AAV30137;

DT 27-OCT-1999 (first entry)

DE Amino acid sequence of tumour suppressor protein NE-dlg.
 KW Nedasin; tumour suppressor protein; NE-dlg; tumour suppression;
 KW malignant tumour formation.

OS Homo sapiens.

PN WO943702-A1.

PD 02-SEP-1999.

PF 24-AUG-1998; 98WO-JP03740.

PR 25-FEB-1998; 98JP-0043552.

PA (SUME) SUMITOMO ELECTRIC IND CO.

PI Kishimoto T, Kuwahara H, Miwa S, Saya H;

DR WPI; 1999-51835/43.

DR N-PSDB; AAX86725.

PT New nedasin protein, useful in the investigation of the mechanism of
 formation of malignant tumours

PS Example 1; Page 81-87; 104pp; Japanese.

CC The present sequence represents tumour suppressor protein NE-dlg. The
 CC specification describes mammalian proteins, designated nedasin, which
 CC bind to the tumour suppressor protein NE-dlg. Nedasin is involved
 CC with the regulation of tumour suppression mechanisms. Nedasin and
 CC antibodies recognizing it are useful in the investigation of the
 CC mechanism of formation of malignant tumours. The polynucleotides are
 CC useful for screening cDNA libraries, and as a source of primers and
 CC probes.

SQ Sequence 817 AA;

Query Match 68.3%; Score 43; DB 20; Length 817;
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRWPTC 8
 |||||

DB 103 PSMWPEC 109

RESULT 15

AAW48102
 ID AAW48102 standard; Protein; 849 AA.

AC AAW48102;

DT 15-JUN-1998 (first entry)

DE Human discs large 1 gene-2 cancer related molecule.

KW Human; NE-dlg; discs large 1 gene; cancer related molecule; nerve;
 internal secretion tissue.

OS Homo sapiens.

PN JP10066581-A.

PD 10-MAR-1998.

PF 23-AUG-1996; 96JP-0241370.

PR 23-AUG-1996; 96JP-0241370.

PA (SUME) SUMITOMO ELECTRIC IND CO.

DR WPI; 1998-224339/20.

DR N-PSDB: AAV20420.

XX Human discs large 1 gene family - useful in, e.g. therapeutic
 PT composition(s) for treating cancer
 XX

PS Claim 8; Page 18-20; 31pp; Japanese.

XX
 CC The present sequence represents human dl9-2 molecule. The present
 CC invention describes human discs large 1 gene (dl9) family expressible in
 CC nerve tissue. Also described are: (I) a polynucleotide (PN) encoding dl9
 CC and comprising a 3100 (I) or 3213 bp (II) sequence; (2) RNA corresponding
 CC to PN having sequences (I) and (II); (3) RNA having a base sequence
 CC translated by dl9; (4) an antisense PN having a at least 15 bp sequence,
 CC and which is a part of PN of (1); (5) a derivative of the antisense PN;
 CC (6) an antibody specific for dl9, and (7) an antibody specific for a
 CC polypeptide having a 817 or 849 aa sequence. The polypeptide, RNA's and
 CC antibodies can be used for detection of dl9. The antisense PN can be
 CC used as a therapeutic composition for treating cancer.

XX
 SQ Sequence 849 AA;

Query Match

68.3%; Score 43; DB 19; Length 849;

Best Local Similarity 71.4%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 PRWMPYC 8

DB 103 PSMWPEC 109

Search completed: March 27, 2003, 16:10:33
 Job time : 6.42105 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 16:08:52 ; Search time 1.94737 Seconds

(without alignments)
135,982 Million cell updates/sec

Title: US-10-019-219-2

Perfect score: 63

Sequence: 1 SPRMPTCL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/iaa/5a_COMB.pep:*

2: /cgn2_6/ptodata/1/iaa/5b_COMB.pep:*

3: /cgn2_6/ptodata/1/iaa/6a_COMB.pep:*

4: /cgn2_6/ptodata/1/iaa/6b_COMB.pep:*

5: /cgn2_6/ptodata/1/iaa/pctus_COMB.pep:*

6: /cgn2_6/ptodata/1/iaa/backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	68.3	49	4	US-09-208-140-20
2	43	68.3	3010	4	US-09-014-416-3
3	40	63.5	520	1	US-08-261-822a-10
4	40	63.5	520	5	PCT-US95-07744A-10
5	39	61.9	376	2	US-08-758-621-10
6	39	61.9	376	4	US-09-107-858-10
7	39	61.9	567	1	US-08-261-822a-12
8	39	61.9	567	5	PCT-US95-07744A-12
9	38	60.3	584	1	US-08-261-822a-8
10	38	60.3	584	5	PCT-US95-07744A-8
11	38	60.3	1182	4	US-09-287-354-6
12	38	60.3	1189	4	US-09-287-354-3
13	38	60.3	1189	4	US-09-287-354-4
14	38	60.3	1207	4	US-09-287-354-5
15	37	58.7	49	3	US-08-851-843a-216
16	37	58.7	49	4	US-08-854-050-216
17	37	58.7	49	4	US-08-854-050-215
18	37	58.7	49	4	US-09-430-323-216
19	37	58.7	49	4	US-08-851-843a-215
20	37	58.7	94	4	US-08-874-549a-334
21	37	58.7	94	4	US-08-854-050-215
22	37	58.7	94	4	US-09-430-323-215
23	37	58.7	311	4	US-09-330-317b-2
24	37	58.7	311	4	US-09-277-716-18
25	37	58.7	311	4	US-09-609-161b-18
26	37	58.7	311	4	US-09-808-589a-2
27	37	58.7	339	4	US-09-330-317b-4

28	37	58.7	339	4	US-09-330-317b-12	Sequence 12, Appl
29	37	58.7	339	4	US-09-330-317b-14	Sequence 14, Appl
30	37	58.7	339	4	US-09-808-589a-4	Sequence 4, Appl
31	37	58.7	339	4	US-09-808-589a-12	Sequence 12, Appl
32	37	58.7	339	4	US-09-808-589a-14	Sequence 14, Appl
33	36	57.1	59	2	US-09-006-675-4	Sequence 4, Appl
34	36	57.1	59	4	US-09-228-603a-4	Sequence 4, Appl
35	36	57.1	104	4	US-09-133-521-2	Sequence 2, Appl
36	36	57.1	333	2	US-08-508-761b-35	Sequence 35, Appl
37	36	57.1	333	3	US-08-997-080-36	Sequence 36, Appl
38	36	57.1	333	2	US-08-997-362-36	Sequence 36, Appl
39	36	57.1	333	3	US-08-873-870-36	Sequence 36, Appl
40	36	57.1	333	4	US-09-095-855-36	Sequence 36, Appl
41	36	57.1	333	4	US-08-705-347a-36	Sequence 36, Appl
42	36	57.1	333	4	US-09-324-542-36	Sequence 36, Appl
43	36	57.1	333	4	US-09-205-426-36	Sequence 36, Appl
44	36	57.1	333	4	US-09-200-643-36	Sequence 36, Appl
45	36	57.1	340	2	US-08-107-676-3	Sequence 3, Appl

ALIGNMENTS

```

RESULT 1
US-09-208-140-20
; Sequence 20, Application US/09208140
; Patent No. 6226576
; GENERAL INFORMATION:
; APPLICANT: Del Vecchio, Alfred
; TITLE OF INVENTION: HEPATITIS C VIRUS NS5B TRUNCATED PROTEIN
; FILE REFERENCE: P50743
; CURRENT APPLICATION NUMBER: US/09/208,140
; CURRENT FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Viral
US-09-208-140-20

Query Match      68.3%; Score 43; DB 4; Length 49;
Best Local Similarity 75.0%; Pred. No. 2.3;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      2 PRMPTCL 9
Db      27 PRMPTCL 34

RESULT 2
US-09-014-416-3
; Sequence 3, Application US/09014416
; Patent No. 6153421
; GENERAL INFORMATION:
; APPLICANT: Yanaqi, Masayuki
; APPLICANT: Bukh, Jens
; APPLICANT: Emerson, Susanne U.
; APPLICANT: Puttelli, Robert H.
; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
; FILE REFERENCE: 20264276
; CURRENT APPLICATION NUMBER: US/09/014,416
; CURRENT FILING DATE: 1998-01-27
; EARLIER APPLICATION NUMBER: US 60/053,062
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3010
; TYPE: PRT
; ORGANISM: Hepatitis C virus

```

US-09-014-416-3

Query Match 68.3%; Score 43; DB 4; Length 3010;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRMWPCL 9
DB 2988 PRMWPCL 2995

RESULT 3

US-08-261-822A-10
; Sequence 10, Application US/08261822A
; Patent No. 5650535
; GENERAL INFORMATION:
; APPLICANT: Ecker, Joseph R. et al.
; TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
; TITLE OF INVENTION: and Pathogens
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553aris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261,822A
; FILING DATE: 17-JUN-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 520 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO

US-08-261-822A-10

Query Match 63.5%; Score 40; DB 1; Length 520;
Best Local Similarity 71.4%; Pred. No. 66;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SPRMWP 7
DB 216 TPRMWP 222

RESULT 4

PCT-US95-07744A-10
; Sequence 10, Application PC/TUS9507744A
; GENERAL INFORMATION:
; APPLICANT: Trustees of The University of Pennsylvania
; TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
; TITLE OF INVENTION: and Pathogens
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
; STREET: One Liberty Place, 46th floor

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07744A
; FILING DATE: 15-JUNE-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,822
; FILING DATE: June 17, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 520 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO

PCT-US95-07744A-10

Query Match 63.5%; Score 40; DB 5; Length 520;
Best Local Similarity 71.4%; Pred. No. 66;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SPRMWP 7
DB 216 TPRMWP 222

RESULT 5

US-08-758-621-10
; Sequence 10, Application US/08758621
; Patent No. 5846821
; GENERAL INFORMATION:
; APPLICANT: Gueriot, Mary Lou, and Elide, David J.
; TITLE OF INVENTION: Metal-Regulated Transporters and Uses Therefor
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/758,621
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/018,578
; FILING DATE: 29-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Silverl, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: DC1-099CP
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-758-621-10

Query Match 61.9%; Score 39; DB 2; Length 376;
Best Local Similarity 66.7%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SPRMWPCTCL 9
Db 277 SKRMWPWAL 285

RESULT 6
US-09-107-858-10
Sequence 10, Application US/09107858
Patent No. 6162900
GENERAL INFORMATION:
APPLICANT: Guerinet, Mary Lou et al.
TITLE OF INVENTION: METAL-REGULATED TRANSPORTERS AND USES THEREFOR
FILE REFERENCE: DCI-099CPDV
CURRENT APPLICATION NUMBER: US/09/107,858
EARLIER FILING DATE: 1998-06-30
CURRENT FILING DATE: 1998-06-30
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 10
LENGTH: 376
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-107-858-10

Query Match 61.9%; Score 39; DB 4; Length 376;
Best Local Similarity 66.7%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SPRMWPCTCL 9
Db 277 SKRMWPWAL 285

RESULT 7
US-08-261-822A-12
Sequence 12, Application US/08261822A
Patent No. 5650553
GENERAL INFORMATION:
APPLICANT: Ecker, Joseph R. et al.
TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
NUMBER OF SEQUENCES: 82
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553-its
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/261,822A
FILING DATE: 17-JUN-1994
CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
NAME: Beardsell, Lori Y.
REGISTRATION NUMBER: 34,293
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 567 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-261-822A-12

Query Match 61.9%; Score 39; DB 1; Length 567;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRMWPCT 7
Db 200 PRMWPCT 205

RESULT 8
PCT-US95-07744A-12
Sequence 12, Application PC/TUS9507744A
GENERAL INFORMATION:
APPLICANT: Trustees of The University of Pennsylvania
TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
NUMBER OF SEQUENCES: 82
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07744A
FILING DATE: 15-JUNE-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/261,822
FILING DATE: June 17, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Beardsell, Lori Y.
REGISTRATION NUMBER: 34,293
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 567 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US95-07744A-12

Query Match 61.9%; Score 39; DB 5; Length 567;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 PRWPT 7
11111
Db 200 PPWPT 205

RESULT 9

US-08-261-822A-8
Sequence 8, Application US/08261822A
Patent No. 5650553
GENERAL INFORMATION:
APPLICANT: Ecker, Joseph R. et al.
TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
NUMBER OF SEQUENCES: 82
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553ris
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/261,822A
FILING DATE: 17-JUN-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 584 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-261-822A-8
Query Match
Best Local Similarity 60.3%; Score 38; DB 1; Length 584;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 SPRWPT 6
11111
Db 213 SPPWPT 218

RESULT 10

PCT-US95-07744A-8
Sequence 8, Application PC/TUS9507744A
GENERAL INFORMATION:
APPLICANT: Trustees of The University of Pennsylvania
TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
NUMBER OF SEQUENCES: 82
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07744A
FILING DATE: 15-JUNE-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/261,822
FILING DATE: June 17, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 584 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
ANTI-SENSE: NO
PCT-US95-07744A-8

Query Match

Best Local Similarity 60.3%; Score 38; DB 5; Length 584;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SPRWPT 6
11111
Db 213 SPPWPT 218

RESULT 11

US-09-287-354-6
Sequence 6, Application US/09287354
Patent No. 6348348
GENERAL INFORMATION:
APPLICANT: THOMPSON, Catherine C.
TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
FILE REFERENCE: Thompson-20263/0243435
CURRENT FILING DATE: 1999-04-07
EARLIER APPLICATION NUMBER: US 60/080,888
FILING DATE: 1998-04-07
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 6
LENGTH: 1182
TYPE: PRT
ORGANISM: Mouse
US-09-287-354-6

Query Match
Best Local Similarity 60.3%; Score 38; DB 4; Length 1182;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SPRWPTCL 9
11111
Db 155 APFWPTCL 163

RESULT 12

US-09-287-354-3
Sequence 3, Application US/09287354
Patent No. 6348348
GENERAL INFORMATION:
APPLICANT: THOMPSON, Catherine C.
TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN

FILE REFERENCE: Thompson-20263/0243435
CURRENT APPLICATION NUMBER: US/09/287/354
CURRENT FILING DATE: 1999-04-07
EARLIER APPLICATION NUMBER: US 60/080,888
EARLIER FILING DATE: 1998-04-07
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 3
LENGTH: 1189
TYPE: PRT
ORGANISM: Homo sapiens
US-09-287-354-3

Query Match 60.3%; Score 38; DB 4; Length 1189;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Caps 0;

QY 1 SPRMWPCTL 9
DB 155 APFWVPCTL 163

RESULT 13
US-09-287-354-4
Sequence 4, Application US/09287354
Patent No. 6348348
GENERAL INFORMATION:
APPLICANT: THOMPSON, Catherine C.
TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
FILE REFERENCE: Thompson-20263/0243435
CURRENT APPLICATION NUMBER: US/09/287/354
CURRENT FILING DATE: 1999-04-07
EARLIER APPLICATION NUMBER: US 60/080,888
EARLIER FILING DATE: 1998-04-07
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 1189
TYPE: PRT
ORGANISM: Homo sapiens
US-09-287-354-4

Query Match 60.3%; Score 38; DB 4; Length 1189;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Caps 0;

QY 1 SPRMWPCTL 9
DB 155 APFWVPCTL 163

RESULT 14
US-09-287-354-5
Sequence 5, Application US/09287354
Patent No. 6348348
GENERAL INFORMATION:
APPLICANT: THOMPSON, Catherine C.
TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
FILE REFERENCE: Thompson-20263/0243435
CURRENT APPLICATION NUMBER: US/09/287/354
CURRENT FILING DATE: 1999-04-07
EARLIER APPLICATION NUMBER: US 60/080,888
EARLIER FILING DATE: 1998-04-07
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 1207
TYPE: PRT
ORGANISM: Rat
US-09-287-354-5

Query Match 60.3%; Score 38; DB 4; Length 1207;
Best Local Similarity 66.7%; Pred. No. 3e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Caps 0;
QY 1 SPRMWPCTL 9
DB 181 APFWVPCTL 189

RESULT 15
US-08-851-843A-216
Sequence 216, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cecch, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 216:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-851-843A-216

Query Match 58.7%; Score 37; DB 3; Length 49;
Best Local Similarity 80.0%; Pred. No. 19;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Caps 0;

QY 4 WMPCT 8
DB 3 WMPCT 7

Mon Apr 7 09:26:00 2003

us-10-019-219-2.rai

Page 6

Search completed: March 27, 2003, 16:13:23
Job time : 2.94737 secs

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 16:12:13 ; Search time 2.15789 Seconds

(without alignments)
244.921 Million cell updates/sec

Title: US-10-019-219-2

Perfect score: 63

Sequence: 1 SPRMWPCL 9

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 237916 segs, 58723674 residues

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCF_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCFUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	73.0	72	10	US-09-764-869-648
2	43	68.3	49	10	US-09-733-183A-20
3	42	66.1	34	10	US-09-864-761-42036
4	41	65.1	68	9	US-09-864-761-42036
5	40	63.5	155	10	US-09-864-761-42036
6	39	61.9	77	10	US-09-864-761-42036
7	38	60.3	1182	9	US-10-024-368-6
8	38	60.3	1189	9	US-10-024-368-3
9	38	60.3	1189	9	US-10-024-368-4
10	38	60.3	1207	9	US-10-024-368-5
11	37.5	59.5	36	9	US-10-050-704-319
12	37.5	59.5	135	9	US-10-050-704-325
13	37.5	59.5	177	9	US-10-050-704-316
14	37.5	59.5	279	10	US-09-815-242-11732
15	37	58.7	17	10	US-09-864-761-46894
16	37	58.7	49	9	US-09-843-676-216
17	37	58.7	49	9	US-09-438-486-216
18	37	58.7	49	9	US-10-053-758-216
19	37	58.7	49	9	US-10-054-293-216

20	37	58.7	94	9	US-09-843-676-215	Sequence 215, App
21	37	58.7	94	9	US-09-438-486-215	Sequence 215, App
22	37	58.7	94	9	US-10-053-758-215	Sequence 215, App
23	37	58.7	94	9	US-10-054-293-215	Sequence 215, App
24	37	58.7	112	9	US-09-764-868-996	Sequence 996, App
25	37	58.7	311	9	US-09-559-874-4	Sequence 4, App11
26	37	58.7	398	9	US-09-965-929-34	Sequence 34, App1
27	37	58.7	1812	10	US-09-775-938A-38	Sequence 1265, App
28	36	57.1	73	10	US-09-925-300-1265	Sequence 42412, A
29	36	57.1	74	10	US-09-864-761-42412	Sequence 4, App11
30	36	57.1	104	9	US-09-920-897-4	Sequence 551, App
31	36	57.1	172	9	US-10-092-154-551	Sequence 551, App
32	36	57.1	172	10	US-09-764-847-551	Sequence 36, App1
33	36	57.1	333	9	US-10-051-643-36	Sequence 36, App1
34	36	57.1	333	9	US-09-880-505-36	Sequence 37, App1
35	36	57.1	340	9	US-10-051-643-37	Sequence 37, App1
36	36	57.1	340	9	US-09-880-505-37	Sequence 14, App1
37	36	57.1	364	10	US-09-051-755-14	Sequence 15, App1
38	36	57.1	364	10	US-09-051-755-15	Sequence 6, App11
39	36	57.1	434	9	US-10-140-372-6	Sequence 5, App11
40	36	57.1	474	10	US-09-887-038-5	Sequence 703, App
41	36	57.1	474	10	US-09-925-297-703	Sequence 326, App
42	35.5	56.3	284	10	US-08-424-550B-326	Sequence 34119, A
43	35	55.6	15	8	US-08-424-550B-326	Sequence 1990, Ap
44	35	55.6	18	10	US-09-864-761-34119	
45	35	55.6	56	10	US-09-867-550-1990	

ALIGNMENTS

RESULT 1
US-09-764-869-648
Sequence 648, Application US/09764869
Patient No. US20020061521A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC007
CURRENT APPLICATION NUMBER: US/09/764, 869
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 2442
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 648
LENGTH: 72
TYPE: PRT
ORGANISM: Homo sapiens
US-09-764-869-648

Query Match
Best Local Similarity 73.0%; Score 46; DB 10; Length 72;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWMPTCL 9
Db 49 RWMPTCL 55

RESULT 2
US-09-733-183A-20
Sequence 20, Application US/09733183A
Patient No. US20020081568A1
GENERAL INFORMATION:
APPLICANT: DelVecchio, Alfred M.
TITLE OF INVENTION: Hepatitis C Virus NS5B Truncated Protein
FILE REFERENCE: P50743D1
CURRENT APPLICATION NUMBER: US/09/733, 183A
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 09/208140
PRIOR FILING DATE: 1998-12-09
PRIOR APPLICATION NUMBER: 60/069208

```
;; PRIOR FILING DATE: 1997-12-11
;; NUMBER OF SEQ ID NOS: 34
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 20
;; LENGTH: 49
;; TYPE: PRT
;; ORGANISM: Viral
US-09-733-183A-20

Query Match
Best Local Similarity 68.3%; Score 43; DB 10; Length 49;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWMPCTC 9
DB 27 PRWMPCTC 34

RESULT 3
US-09-864-761-42036
; Sequence 42036, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42036
; LENGTH: 34
; TYPE: PRT
```

```
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC017004.2
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.3
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.3
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.5
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 12
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.9
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.8
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 12
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 14
US-09-864-761-42036

Query Match
Best Local Similarity 66.7%; Score 42; DB 10; Length 34;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWMPTC 8
DB 19 RWMPTC 24

RESULT 4
US-09-986-480-449
; Sequence 449, Application US/0986480
; Publication No. US20030027999A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 143 Human Secreted Proteins
; FILE REFERENCE: P5500P1
; CURRENT APPLICATION NUMBER: US/09/986,480
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/12788
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/134,068
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 456
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 449
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-986-480-449

Query Match
Best Local Similarity 65.1%; Score 41; DB 9; Length 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWMP 6
DB 6 PRWMP 10

RESULT 5
US-09-729-674-104
; Sequence 104, Application US/09729674
; Patent No. US20010039335A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racle, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Metberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
```

APPLICANT: Wong, Gordon G.
 APPLICANT: Clark, Hilary
 APPLICANT: Fechtel, Kim
 APPLICANT: Genetics Institute, Inc.
 TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 FILE REFERENCE: 6055-64X
 CURRENT APPLICATION NUMBER: US/09/729,674
 CURRENT FILING DATE: 2000-12-04
 PRIOR APPLICATION NUMBER: 09/539,330
 PRIOR FILING DATE: 2000-03-30
 NUMBER OF SEQ ID NOS: 283
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 104
 LENGTH: 155
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: UNSURE
 LOCATION: (46)
 US-09-729-674-104

Query Match 63.5%; Score 40; DB 10; Length 155;
 Best Local Similarity 62.5%; Pred. No. 36;
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PRWMPCT 9
 DB 61 PAMWACL 68

RESULT 6
 US-09-864-761-36434
 Sequence 36434, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharron G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 FILE REFERENCE: Aecmics-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 36434
 LENGTH: 77
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC008018.18
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
 OTHER INFORMATION: EST_HUMAN HIT: BE965190.2, EVALUATE 1.20e-01
 OTHER INFORMATION: SWISSPROT HIT: P02919, EVALUATE 2.60e+00
 US-09-864-761-36434

Query Match 61.3%; Score 39; DB 10; Length 77;
 Best Local Similarity 62.5%; Pred. No. 28;
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SPRWMPCT 8
 DB 37 SPLMWEVC 44

RESULT 7
 US-10-024-368-6
 Sequence 6, Application US/10024368
 Publication No. US20030027300A1
 GENERAL INFORMATION:
 APPLICANT: THOMPSON, Catherine C.
 TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
 FILE REFERENCE: Thompson-20263/0243435
 CURRENT APPLICATION NUMBER: US/10/024,368
 CURRENT FILING DATE: 2001-12-21
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
 PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
 PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 6
 LENGTH: 1182
 TYPE: PRT
 ORGANISM: Mouse
 US-10-024-368-6

Query Match 60.3%; Score 38; DB 9; Length 1182;
 Best Local Similarity 66.7%; Pred. No. 4.2e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPRWMPCT 9
 DB 155 APFWMPCT 163

RESULT 8
 US-10-024-368-3
 Sequence 3, Application US/10024368
 Publication No. US20030027300A1

```
;; GENERAL INFORMATION:
;; APPLICANT: THOMPSON, Catherine C.
;; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
;; FILE REFERENCE: Thompson-20263/0243435
;; CURRENT APPLICATION NUMBER: US/10/024,368
;; CURRENT FILING DATE: 2001-12-21
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
;; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
;; NUMBER OF SEQ ID NOS: 10
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 3
;; LENGTH: 1189
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-024-368-3

Query Match      60.3%; Score 38; DB 9; Length 1189;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY      1 SPRWMPCTL 9
        :| | | | |
Db      155 APFWPCTL 163

RESULT 9
US-10-024-368-4
; Sequence 4, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-368-4

Query Match      60.3%; Score 38; DB 9; Length 1189;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY      1 SPRWMPCTL 9
        :| | | | |
Db      155 APFWPCTL 163

RESULT 10
US-10-024-368-5
; Sequence 5, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
```

```
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 5
;; LENGTH: 1207
;; TYPE: PRT
;; ORGANISM: Rat
US-10-024-368-5

Query Match      60.3%; Score 38; DB 9; Length 1207;
Best Local Similarity 66.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY      1 SPRWMPCTL 9
        :| | | | |
Db      181 APFWPCTL 189

RESULT 11
US-10-050-704-319
; Sequence 319, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 319
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-050-704-319

Query Match      59.5%; Score 37.5; DB 9; Length 36;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

OY      2 PRW---WPTC 8
        || | | | |
Db      5 PRWTWMPIC 14

RESULT 12
US-10-050-704-325
; Sequence 325, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 325
; LENGTH: 135
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-050-704-325

```

```

Query Match
Best Local Similarity 59.5%; Score 37.5; DB 9; Length 135;
Matches 6; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

```

```

QY 2 PRW---WPTC 8
   ||| |||
Db 34 PRWTWLPIC 43

```

```

RESULT 13
US-10-050-704-316
; Sequence 316, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 316
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-050-704-316

```

```

Query Match
Best Local Similarity 59.5%; Score 37.5; DB 9; Length 177;
Matches 6; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

```

```

QY 2 PRW---WPTC 8
   ||| |||
Db 76 PRWTWLPIC 85

```

```

RESULT 14
US-09-815-242-11732
; Sequence 11732, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578

```

```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11732
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-815-242-11732

```

```

Query Match
Best Local Similarity 59.5%; Score 37.5; DB 10; Length 279;
Matches 7; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

```

```

QY 1 SPRW---WPTCL 9
   ||| |||
Db 17 SPWWRMCWPCCL 28

```

```

RESULT 15
US-09-864-761-46894
; Sequence 46894, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenshang
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203

```

; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 46894
 ; LENGTH: 17
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AL118499.5
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96
 US-09-864-761-46894

Query Match 58.7%; Score 37; DB 10; Length 17;
 Best Local Similarity 80.0%; Pred. No. 14;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 WMPYC 8
 DB 4 WMPYC 8

Search completed: March 27, 2003, 16:20:26
 Job time : 3.15789 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 16:01:27 ; Search time 1.52632 Seconds

(without alignments)
566,862 Million cell updates/sec

Title: US-10-019-219-2

Perfect score: 63

Sequence: 1 SPRMWPCTL 9

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	73.0	806	2	G84863
2	43	68.3	120	2	T42054
3	42	66.7	221	2	C81099
4	42	66.7	221	2	C81842
5	42	66.7	275	2	I40211
6	42	66.7	443	2	I39538
7	42	66.7	2092	2	S30026
8	42	66.7	2149	2	S18676
9	41	65.1	123	2	H72698
10	41	65.1	252	1	J00417
11	41	65.1	235	2	AD3577
12	41	65.1	461	2	H70899
13	41	65.1	790	2	T34293
14	41	65.1	1398	2	T18350
15	40	63.5	64	2	S28486
16	40	63.5	212	2	S74376
17	40	63.5	263	1	S43189
18	40	63.5	290	1	D47468
19	40	63.5	471	2	T50016
20	40	63.5	542	2	I39540
21	40	63.5	997	2	T39521
22	39	61.9	119	2	A53257
23	39	61.9	176	2	H75332
24	39	61.9	338	2	A87566
25	39	61.9	376	2	S33654
26	38.5	61.1	567	2	E96764
27	38.5	61.1	423	2	AG2394
28	38	60.3	170	2	S26718
29	38	60.3	175	2	T27543

30	38	60.3	197	2	S59397	probable membrane
31	38	60.3	211	1	MUKAD	lysosome (EC 3.2.1
32	38	60.3	248	2	S23449	NADH oxidase (H2O2
33	38	60.3	335	2	T46351	hypothetical prote
34	38	60.3	369	2	T40279	hypothetical prote
35	38	60.3	374	2	T43708	cytochrome oxidase
36	38	60.3	379	2	E83597	homoserine O-acety
37	38	60.3	528	2	B75310	conserved hypothet
38	38	60.3	584	2	B84668	ethylene-insensiti
39	38	60.3	598	2	C82194	ATP-dependent prot
40	38	60.3	731	2	JC7701	ARHGAP9 protein -
41	38	60.3	770	2	C87316	alpha-N-acetylgluc
42	38	60.3	1162	2	I48378	hairless protein -
43	37	58.7	133	2	F72471	hypothetical prote
44	37	58.7	149	2	A87346	hypothetical prote
45	37	58.7	247	2	A70910	probable truncated

ALIGNMENTS

RESULT 1

G84863 hypothetical protein At2g43240 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: G84863

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKien, S.E.; Umayam, L.; Tallon,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: G84863

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-806 <STO>

A:Cross-references: GB:AE002093; MID:g3763933; PIDN:AAC64313.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g43240

A:Map position: 2

Query Match 73.0%; Score 46; DB 2; Length 806;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRMWPCTL 9

DB 676 PKMWPCTL 683

RESULT 2

T42054 hypothetical protein - Streptomyces coelicolor (fragment)

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T42054

R:Guillarro, J.; Santamaría, R.; Schauer, A.; Losick, R.

J. Bacteriol. 170, 1895-1901, 1988

A:Title: Promoter determining the timing and spatial localization of transcription of

A:Reference number: Z22043; MUID:88169521; PMID:2450872

A:Accession: T42054

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-120 <GUI>

A:Cross-references: EMBL:M20145; PIDN:AAA26812.1

Query Match 68.3%; Score 43; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WMPWC 8

IIIIII

DB 88 RWPTC 92

RESULT 3

C81099

hypothetical protein NMB1296 [imported] - *Neisseria meningitidis* (strain MC58 serogroup C)
 C:Species: *Neisseria meningitidis*
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: C81099
 R:Retelien, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
 Li, H.; Qin, H.; Vamathevan, J.; Gall, J.; Scarlato, V.; Maignan, V.; Pizza, M.
 Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
 A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
 A:Reference number: AB1000; MUID:20175755; PMID:10710307

A:Accession: C81099
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-221 <TEP>
 A:Cross-references: GB:AE002478; GB:AE002098; NID:g7226533; PIDN:AAFA1672.1; PID:g722653

A:Experimental source: serogroup B, strain MC58
 C:Genetics:

A:Gene: NMB1296

Query Match 66.7%; Score 42; DB 2; Length 221;
 Best Local Similarity 83.3%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 RWMPIC 8
 DB 177 RWWPQC 182

RESULT 4
 C81842

hypothetical protein NMA1506 [imported] - *Neisseria meningitidis* (strain Z2491 serogroup C)
 C:Species: *Neisseria meningitidis*
 C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C:Accession: C81842
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

; Holroyd, S.; Jorgels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
 A:Reference number: AB1775; MUID:2022556; PMID:10761919

A:Accession: C81842
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-221 <PAR>
 A:Cross-references: GB:AL162756; GB:AL157959; NID:g7380091; PIDN:CAB84738.1; PID:g738015

A:Experimental source: serogroup A, strain Z2491
 C:Genetics:

A:Gene: NMA1506

Query Match 66.7%; Score 42; DB 2; Length 221;
 Best Local Similarity 83.3%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 RWMPIC 8
 DB 177 RWWPQC 182

RESULT 5
 I40211

probable sterol dehydrogenase (EC 1.1.1.-) - *Bradyrhizobium japonicum*
 C:Species: *Bradyrhizobium japonicum*
 C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 08-Oct-1999

C:Accession: I40211
 R:Tully, R.E.; Keister, D.L.
 Appl. Environ. Microbiol. 59, 4136-4142, 1993

A:Title: Cloning and mutagenesis of a cytochrome P-450 locus from *Bradyrhizobium japonicum*
 A:Reference number: I40207

Query Match 66.7%; Score 42; DB 2; Length 221;
 Best Local Similarity 83.3%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 RWMPIC 8
 DB 177 RWWPQC 182

RESULT 6
 I19538

alpha-amylase - *Aeromonas hydrophila*
 C:Species: *Aeromonas hydrophila*
 C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 03-Dec-1999

C:Accession: I19538
 R:Chang, M.C.; Chang, J.C.; Chen, J.P.
 J. Gen. Microbiol. 139, 3215-3223, 1993

A:Title: Cloning and nucleotide sequence of an extracellular alpha-amylase gene from
 A:Reference number: I19538; MUID:94172314; PMID:8126440

A:Accession: I19538
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-443 <RES>
 A:Cross-references: GB:L19299; NID:g304014; PIDN:AAA21016.1; PID:g304015

C:Genetics:
 A:Gene: amya
 A:Superfamily: mammalian alpha-amylase; alpha-amylase core homology
 F:166-291/Domain: alpha-amylase core homology <AMT>

Query Match 66.7%; Score 42; DB 2; Length 443;
 Best Local Similarity 83.3%; Pred. No. 39;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 RWMPIC 8
 DB 406 RWWPQC 411

A:Accession: I40211

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-275 <RES>

A:Cross-references: EMBL:U12678; NID:g529961; PIDN:AAC28892.1; PID:g529965

C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

C:Keywords: oxidoreductase

F:7-190/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 66.7%; Score 42; DB 2; Length 275;

Best Local Similarity 71.4%; Pred. No. 25;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 PRWMPIC 8

DB 219 PRWMPIC 225

RESULT 7

S30026

genome polypeptide - Rift Valley fever virus

N:Contans: RNA-directed RNA polymerase (EC 2.7.7.48)

C:Species: Rift Valley fever virus

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999

C:Accession: S30026

R:Miller, R.; Argentin, C.; Bouloy, M.; Prehaud, C.; Bishop, D.H.L.

Nucleic Acids Res. 20, 6440, 1992

A:Title: Corrigendum: Completion of the genome sequence of Rift Valley fever phlebov

A:Reference number: S30026

A:Accession: S30026

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: genomic RNA

A:Residues: 1-2092 <MUL>

A:Cross-references: GB:X56464

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1992

C:Superfamily: Uukuniemi virus RNA-directed RNA polymerase

C:Keywords: nucleotidyltransferase

Query Match 66.7%; Score 42; DB 2; Length 2092;

Best Local Similarity 83.3%; Pred. No. 1.5e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SPRWMP 6

DB 111111

Db 1013 SPRMWP 1018

RESULT 8
S18676

genome polypeptide - Rift Valley fever virus
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: Rift Valley fever virus
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 24-Nov-1999
C:Accession: S18676
R:Miller, R.; Argentin, C.; Bouliou, M.; Prehaud, C.; Bishop, D.H.L.
Nucleic Acids Res. 19, 5433, 1991
A:Title: Completion of the genome sequence of Rift Valley fever phlebovirus indicates th
A:Reference number: S18676; M01D:92020238; PMID:1923828
A:Accession: S18676
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: genomic RNA
A:Residues: 1-2149 <M01D>
A:Cross-references: EMBL:X56464; NID:g61926; PIDN:CA93836.1; PID:g61927
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, August 1991
C:Superfamily: Tumorvirus RNA-directed RNA polymerase
C:Keywords: nucleotidyltransferase

Query Match 66.7%; Score 42; DB 2; Length 2149;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SPRMWP 6
Db 1013 SPRMWP 1018RESULT 9
H72698

hypothetical protein APE1008 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takai
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyperthermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; M01D:99310339; PMID:10382966
A:Accession: H72698
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-123 <KAW>
A:Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BA79992.1; PID:d1043778; PID:g510
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1008
C:Superfamily: Aeropyrum pernix hypothetical protein APE1008

Query Match 65.1%; Score 41; DB 2; Length 123;
Best Local Similarity 85.7%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 RMMPTCL 9
Db 110 RMMPTCL 116RESULT 10
J00417

oleoyl-[acyl-carrier-protein] hydrolase (EC 3.1.2.14) homolog - *Vibrio anguillarum*
N:Alternate names: hypothetical 28.1k protein; ORF6 protein; S-acyl fatty acid synthase
C:Species: *Vibrio anguillarum*
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: J00417; S26422
R:Farrell, D.H.; Mikesell, P.; Actis, L.A.; Crosa, J.H.
Gene 86, 45-51, 1990
A:Title: A regulatory gene, *angr*, of the iron uptake system of *Vibrio anguillarum*: simil
A:Reference number: J00416; M01D:90185247; PMID:2311935

A:Accession: J00417

A:Molecule type: DNA
A:Residues: 1-252 <FAR>
A:Cross-references: GB:34504; NID:g155150; PIDN:AA79861.1; PID:g155153
R:Polmasky, M.E.; Actis, L.A.; Malsbender, L.S.; Crosa, J.H.
submitted to the EMBL Data Library, April 1992
A:Description: Genetic characterization of the regulatory protein *angr*: presence of 1
A:Reference number: S26421
A:Accession: S26422
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-252 <TOL>
A:Cross-references: EMBL:Z12000; NID:g48322; PIDN:CA78045.1; PID:g48324
C:Comment: This pathogenic bacterium is a causative agent of vibriosis, a widespread
C:Superfamily: oleoyl-[acyl-carrier-protein] hydrolase
C:Keywords: thioester hydrolase
F:19-234/Domains: oleoyl-[acyl-carrier-protein] hydrolase homology <ACPR>

Query Match 65.1%; Score 41; DB 1; Length 252;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SPRMPTCL 9
Db 156 SPRMPTCL 164RESULT 11
AD3577

sugar transport system permease protein [imported] - *Brucella melitensis* (strain 16M)
C:Species: *Brucella melitensis*
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
R:DelVecchio, V.G.; Kaptehal, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanov
; Mazur, M.; Goldstein, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melit*
A:Reference number: AD3577; PMID:11756688
A:Accession: AD3577
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-295 <KRW>
A:Cross-references: GB:AE008918; PIDN:AAU53783.1; PID:g17984712; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0541
A:Map position: II
C:Superfamily: maltose transport protein malG

Query Match 65.1%; Score 41; DB 2; Length 295;
Best Local Similarity 71.4%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SPRMPT 7
Db 63 SPRMPT 69RESULT 12
H70899

probable cytochrome P450 *Pv1394c* [similarity] - *Mycobacterium tuberculosis* (strain H3
C:Species: *Mycobacterium tuberculosis*
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 28-Jul-2000
C:Accession: H70899
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garner, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete geno
A:Reference number: A70500; M01D:98295987; PMID:9634220
A:Accession: H70899
A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-461 <COL>
A:Cross-references: GB:AL123456; NID:g3256012; PIDN:CA802176.1; PID:g1542902
A:Experimental source: strain H37Rv
C:Genetics:

A:Gene: RV1394c
C:Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
F:271-431/Domain: cytochrome P450 homology <P45>
F:409/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 65.1%; Score 41; DB 2; Length 461;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 PRMWP 6
DB 373 PRMWP 377

RESULT 13

T34293
hypothetical protein F49E10.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
C:Accession: T34293

R:Miller, N.

submitted to the EMBL Data Library, April 1996

A:Description: The sequence of C. elegans cosmid F49E10.

A:Reference number: 221500

A:Accession: T34293

A>Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-790 <MIL>

A:Cross-references: EMBL:U53341; PIDN:AAC69106.1; GSPDB:GN00028; CESP:F49E10.2

A:Experimental source: strain Bristol N2; clone F49E10

C:Genetics:

A:Gene: CESP:F49E10.2

A:Map position: X

A:introns: 51/1; 92/3; 176/3; 235/3; 332/2; 514/1; 543/2; 569/3; 677/1; 732/3

Query Match 65.1%; Score 41; DB 2; Length 790;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 PRMWP 6
DB 688 PRMWP 692

RESULT 14

T18350
probable pol polypeptide - rice blast fungus gypsy retroelement (fragment)

C:Species: Magnaporthe grisea (rice blast fungus)

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18350

R:Dodinson, K.F.

submitted to the EMBL Data Library, September 1994

A:Description: Sequence of the grh retroelement.

A:Reference number: 218883

A:Accession: T18350

A>Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1398 <DOB>

A:Cross-references: EMBL:M77661; NID:g538065; PID:g538067; PIDN:AAA21442.1

C:Genetics:

A:Mobile element: gypsy retroelement

Query Match 65.1%; Score 41; DB 2; Length 1398;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 PRMWP 6

DB 76 PRMWP 80

RESULT 15

S28486
hypothetical protein 2 - Vibrio cholerae

C:Species: Vibrio cholerae

C>Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 08-Oct-1999
C:Accession: S28486

R:Manning, P.A.

submitted to the EMBL Data Library, May 1991

A:Reference number: S28467

A:Accession: S28486

A:Molecule type: DNA

A:Residues: 1-64 <MAN>

A:Cross-references: EMBL:X59554; NID:g48381; PIDN:CAA42152.1; PID:g48401

A:Experimental source: strain 017

Query Match 63.5%; Score 40; DB 2; Length 64;
Best Local Similarity 75.0%; Pred. No. 14;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SPRMPTC 8
DB 51 SYRMPTC 58

Search completed: March 27, 2003, 16:12:41
Job time : 3.52632 secs

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 16:00:22 ; Search time 1.52632 Seconds

(without alignments)
244.567 Million cell updates/sec

Title: US-10-019-219-2

Perfect score: 63

Sequence: 1 SPRMPTCL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	68.3	817	1	DLG3_HUMAN
2	43	68.3	849	1	DLG3_RAT
3	42	66.7	275	1	YCP1_BRAJA
4	42	66.7	443	1	AMYA_AERHY
5	42	66.7	849	1	DLG3_MOUSE
6	42	66.7	2149	1	RRPL_RVEFV
7	41	65.1	252	1	SASF_VIBAN
8	41	65.1	461	1	CL32_MYCTU
9	39	61.9	119	1	ATP6_NAEFO
10	39	61.9	376	1	ZRT1_YEAST
11	38	60.3	175	1	YCCA_CAEEL
12	38	60.3	211	1	LYCH_CHASP
13	38	60.3	374	1	OX11_SCHPO
14	38	60.3	379	1	MERY_PSEAE
15	38	60.3	417	1	OBSE_DROME
16	38	60.3	1181	1	HAIR_RAT
17	38	60.3	1182	1	HAIR_MOUSE
18	38	60.3	1189	1	HAIR_HUMAN
19	37	58.7	309	1	ER25_YEAST
20	37	58.7	311	1	LUCT_YENRE
21	37	58.7	418	1	VLAR_SHEEP
22	37	58.7	460	1	YS85_MYCTU
23	37	58.7	546	1	LNT_TREPA
24	37	58.7	809	1	YARA_SCHPO
25	37	58.7	1131	1	YANC_SCHPO
26	37	58.7	1385	1	YMS_CAEEL
27	36.5	57.9	830	1	HMT1_SCHPO
28	36	57.1	94	1	YORE_NMV
29	36	57.1	283	1	ISPE_CHLAM
30	36	57.1	333	1	AB5C_MYCLE
31	36	57.1	340	1	AB5C_MYCTU
32	36	57.1	352	1	AB5C_MYCAV
33	36	57.1	431	1	CIT1_ECOLI

34	36	57.1	434	1	CIT1_SALTY
35	36	57.1	696	1	YIR8_YEAST
36	36	57.1	790	1	ATSY_SYN7
37	36	57.1	1026	1	BGAL_STRTR
38	36	57.1	1165	1	POL_GALV
39	36	57.1	1398	1	POLG_BDVN
40	35.5	56.3	217	1	GRB2_XENLA
41	35	55.6	136	1	YCV4_YEAST
42	35	55.6	236	1	LIPB_MYCPU
43	35	55.6	246	1	MTGA_HAEIN
44	35	55.6	335	1	LPKX_CAUCR
45	35	55.6	351	1	YKX2_CAEEL
					P24115 salmonella
					P40483 saccharomyc
					P37385 synchococc
					P23689 streptococc
					P21414 gibbon ape
					P19711 bovine vira
					P87379 xenopus lae
					P25640 saccharomyc
					Q918P6 mycoplasma
					P44890 haemophilus
					P58184 cauliobacter
					P34327 caenorhabdi

ALIGNMENTS

RESULT 1
ID DLG3_HUMAN STANDARD; PRT; 817 AA.
AC Q92796; Q9ULI8;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Presynaptic protein SAP102 (Synapse-associated protein 102)
DE (Neuroendocrine-DLG) (NE-DLG) (Discs, large homolog 3).
GN DLG3 OR KIAA1232.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
ON NCBI_TaxID=9606;
RX MEDLINE=97333623; PubMed=9188857;
RA Makino K., Kuwahara H., Masuko N., Nishiyama Y., Morisaki T.,
SA Sasaki J., Nakao M., Kuwano A., Nakata M., Ushio Y., Saza H.;
RT Cloning and characterization of NE-dlg: a novel human homolog of the
RT Drosophila discs large (dlg) tumor suppressor protein interacts with
RT the APC protein.";
RL Oncogene 14:2425-2433(1997).
RN [2]
RP SEQUENCE OF 330-817 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hirosewa M., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
CC -!- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF THE NMDA RECEPTOR
SUBUNIT NR2B (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 GUANYLATE KINASE-LIKE DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation on
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U49089; AAB61453.1; -;
DR EMBL: AB033058; BAA86546.1; -;
DR HSSP: Q12959; 1PDR.
DR Genew: HGNC:2902; DLG3.
DR MIM: 300189; -;
DR InterPro: IPR000619; Guanylate_kin.
DR InterPro: IPR001478; PDZ.

DR InterPro: IPR001452; SH3.
DR Pfam: PF00018; SH3; 1.
DR Pfam: PF00595; PDZ; 3.
DR Pfam: PF00625; Guanylate_kin; 1.
DR ProDom: PD000066; SH3; 1.
DR SMART: SM00228; GUKC; 1.
DR SMART: SM00228; PDZ; 3.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
DR PROSITE: PS50106; PDZ; 3.
DR PROSITE: PS50002; SH3; 1.
DR SH3 domain; Repeat.
KW SH3 domain; Repeat.
FT DOMAIN 130 217 PDZ 1.
FT DOMAIN 226 311 PDZ 2.
FT DOMAIN 379 465 PDZ 3.
FT DOMAIN 503 568 SH3.
FT DOMAIN 628 803 GUANYLATE KINASE.
FT CONFLICT 330 381 FTALADNHSHNSLIGTIGAVESKYSPAPQVPTRYSPI
PRHMLAEDFT -> AARREGAMERARKFGSGIAMLIGS
ASASNRASQRMWPLSLRPGDA (IN REF. 2).
DEPGISLDYGAKNL -> SIKTRKKSFRLSRKPFYKSK
ENMAOESSIOEGVTSNTSDSESS (IN REF. 2).
SQ SEQUENCE 817 AA; 90344 MW; 3D7512EC4713FC4E CRC64;
FT CONFLICT 592 606
FT Query Match 68.3%; Score 43; DB 1; Length 817;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 2 PRMWPTC 8
DB 103 PSMWPEC 109
ID DLG3_RAT STANDARD: PRT; 849 AA.
AC Q62936; P70547;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Presynaptic protein SAP102 (Synapse-associated protein 102) (PSD-
95/SAP90 related protein 1) (Discs, large homolog 3).
GN DLG3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (LONG FORM).
RC TISSUE=Brain;
RX MEDLINE=96374358; PubMed=8780649;
RA Mueller B.M., Kistner U., Chung W.J., Kuhlendahl S.,
RA Fenster S.D., Lau L.-F., Voh R.W., Huganir R.L., Gundelfinger E.D.,
RA Garner C.C.;
RT "SAP102, a novel postsynaptic protein that interacts with NMDA
RT receptor complexes in vivo.";
RL Neuron 17:255-265(1996).
[2]
RP SEQUENCE FROM N.A. (SHORT FORM).
RA Irlie M., Hata Y., Takai Y.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF THE NMDA RECEPTOR
SUBUNIT NR2B.
CC -1- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U50147; AAA93031.1; -
DR EMBL: U53367; AAB48561.1; -
DR HSSP: Q12959; IPDR.
DR InterPro: IPR000619; Guanylate_kin.
DR InterPro: IPR001478; PDZ.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00018; SH3; 1.
DR Pfam: PF00595; PDZ; 3.
DR Pfam: PF00625; Guanylate_kin; 1.
DR ProDom: PD000066; SH3; 1.
DR SMART: SM00072; GUKC; 1.
DR SMART: SM00228; PDZ; 3.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
DR PROSITE: PS50106; PDZ; 3.
DR PROSITE: PS50002; SH3; 1.
DR SH3 domain; Repeat; Alternative splicing.
KW SH3 domain; Repeat.
FT DOMAIN 149 235 PDZ 1.
FT DOMAIN 244 330 PDZ 2.
FT DOMAIN 404 484 PDZ 3.
FT DOMAIN 519 589 SH3.
FT DOMAIN 659 849 GUANYLATE KINASE.
FT VARSPLIC 627 640 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 849 AA; 93539 MW; 34DA9C46C7B96DB CRC64;
FT CONFLICT 592 606
FT Query Match 68.3%; Score 43; DB 1; Length 849;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 2 PRMWPTC 8
DB 103 PSMWPEC 109
ID YCPL_BRAJA STANDARD: PRT; 275 AA.
AC Q45219;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Probable short-chain type dehydrogenase/reductase (EC 1.-.-.-).
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=98322110; PubMed=9655913;
RA Tully R.E., van Berkum P., Lovins K.W., Keister D.L.;
RT "Cloning and mutagenesis of a cytochrome P-450 locus from
RT Bradyrhizobium japonicum that is expressed anaerobically and
RT symbiotically.";
RL Appl. Environ. Microbiol. 59:4136-4142(1993).
[2]
RP SEQUENCE FROM N.A.
RA STRAIN=USDA 110;
RX MEDLINE=98322110; PubMed=9655913;
RA Tully R.E., van Berkum P., Lovins K.W., Keister D.L.;
RT "Identification and sequencing of a cytochrome P450 gene cluster from
RT Bradyrhizobium japonicum.";
RL Biochim. Biophys. Acta 1398:243-255(1998).
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
(SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U12678; AAC28892.1; -
DR HSSP: P08074; 1CYP
DR InterPro: IPR002198; ADH_short.
DR Pfam: PF00106; adh_short; 1.
DR PRINTS: PR00080; SDRFAMILY.
DR PROSITE: PS00061; ADH_SHORT; 1.
KW Hypothetical protein: Oxidoreductase.
FT NP_BIND 10 NAD OR NADP (BY SIMILARITY).
FT ACT_SITE 159 159 BY SIMILARITY.
SQ SEQUENCE 275 AA; 28827 MW; 312E7070C404DF86 CRC64;

QY 2 PRMPTC 8
Db 219 PRMWRSC 225
Query Match 66.7%; Score 42; DB 1; Length 275;
Best Local Similarity 71.4%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

RESULT 4
AMYA_AERHY
ID AMYA_AERHY STANDARD; PRT; 443 AA.
AC P41131;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase).
GN AMYA.
OS Aeromonas hydrophila.
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
OC Aeromonas.
OX NCBI_TaxID=644;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MCC-1;
RX MEDLINE=94172314; PubMed=8126440;
RA Chang M.C., Chang J.C., Chen J.P.;
RT Cloning and nucleotide sequence of an extracellular alpha-amylase
RT gene from Aeromonas hydrophila MCC-1.";
RL J. Gen. Microbiol. 139:3215-3223(1993).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U19299; AAC21016.1; -
DR HSSP: P29957; 1AOM.
DR InterPro: IPR000461; Alpha_amylase.
DR Pfam: PF00128; alpha-amylase; 1.
DR Pfam: PF02806; alpha-amylase_C; 1.
DR PRINTS: PR0110; ALPHAAMYLASE.
KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Signal.
FT SIGNAL 1 24
FT CHAIN 25 443 ALPHA-AMYLASE.
FT ACT_SITE 198 198 BY SIMILARITY.
FT ACT_SITE 202 202 BY SIMILARITY.
FT ACT_SITE 287 287 BY SIMILARITY.

SQ SEQUENCE 443 AA; 48333 MW; 8F8D60B9341A92P9 CRC64;

QY 3 PRMPTC 8
Db 406 PRMWRSC 411
Query Match 66.7%; Score 42; DB 1; Length 443;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 5
DLG3_MOUSE
ID DLG3_MOUSE STANDARD; PRT; 849 AA.
AC P70175;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Presynaptic protein SAP102 (Synapse-associated protein 102) (Disca,
DE large homolog 3).
GN DLG3 OR DLG3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6; TISSUE=Brain;
RA Kohmura N., Makino S., Yagi T.;
RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INTERACTS WITH THE CYTOSOLIC TAIL OF THE NMDA RECEPTOR
CC SUBUNIT NR2B.
CC -1- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D87117; BAA13249.1; -
DR HSSP: Q12959; 1PDR.
DR MGD: MGI:1888986; DLG3.
DR InterPro: IPR000619; Guanylate_kin.
DR InterPro: IPR001478; PDZ.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00018; SH3; 1.
DR Pfam: PF00595; PDZ; 3.
DR Pfam: PF00625; Guanylate_kin; 1.
DR ProDom: PD000066; SH3; 1.
DR SMART: SM00072; GUKC; 1.
DR SMART: SM00228; PDZ; 3.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
DR PROSITE: PS50106; PDZ; 3.
DR PROSITE: PS50002; SH3; 1.
KW SH3 domain; Repeat.
FT DOMAIN 149 235 PDZ 1.
FT DOMAIN 244 330 PDZ 2.
FT DOMAIN 404 484 PDZ 3.
FT DOMAIN 519 589 SH3.
FT DOMAIN 659 849 GUANYLATE_KINASE.
SQ SEQUENCE 849 AA; 93482 MW; EF3EF2D7513538EE CRC64;

QY 3 PRMPTC 8
Db 406 PRMWRSC 411
Query Match 66.7%; Score 42; DB 1; Length 849;
Best Local Similarity 71.4%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRMPTC 8
11111
Db 103 PGWPEC 109

RESULT 6

RRPL_RVFVZ STANDARD: PRT: 2149 AA.

AC P27316;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA-directed RNA polymerase (PC 2.7.7.48) (L protein).

GN L.
OS Rift valley fever virus (strain ZH-548 M12) (RVFV).

OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.

OX NCBI_TaxID=11589;

RP SEQUENCE FROM N.A.

RX MEDLINE=92020238; PubMed=1923828;

RA Mueller R., Argenti C., Bouloy M., Prehaud C., Bishop D.H.L.;

RT "Completion of the genome sequence of Rift Valley fever phlebovirus indicates that the L RNA is negative sense and codes for a putative

transcriptase-replicase.";

RL Nucleic Acids Res. 19:5433-5433(1991).

CC -1 CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +

CC [RNA](N).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR EMBL: X56464; CAA39836.1; -

DR PIR: S18676; S18676.

RW TRANSFERASE; Nucleotidyltransferase; RNA-directed RNA polymerase.

SQ SEQUENCE 2149 AA; 243589 MW; 8D5739C6079A8BD7 CRC64;

Query Match 66.7%; Score 42; DB 1; Length 2149;
Best Local Similarity 83.3%; Pred. No. 71;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPRMWP 6
11111
Db 1013 SPRMWP 1018

RESULT 7
SAST_VIBAN STANDARD: PRT: 252 AA.

ID SAST_VIBAN STANDARD: PRT: 252 AA.

AC P19829;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Probable anguillabactin biosynthesis thioesterase (EC 3.1.2.-).

OS Vibrio anguillarum (Listonella anguillarum).

OG Plasmid pM1.

OC Bacteria; Proteobacteria; gamma subdivision: Vibrionaceae; Listonella.

OX NCBI_TaxID=35601;

RP SEQUENCE FROM N.A.

RX MEDLINE=90185247; PubMed=2311935;

RA Farrell D.H., Mikesell P., Actis L.A., Crosa J.H.;

RT "A regulatory gene, *angr*, of the iron uptake system of *Vibrio anguillarum*: similarity with phage P22 *cro* and regulation by iron.";

STRAIN-531A;

MEDLINE=93328275; PubMed=8335354;

RA Tolmasey M.E., Actis L.A., Crosa J.H.;

RT "A single amino acid change in *angr*, a protein encoded by pM1-1 like

virulence plasmids, results in hyperproduction of anguillabactin.";

RL Infect. Immun. 61:3228-3233(1993).

CC -1 FUNCTION: PROBABLE THIOESTERASE INVOLVED IN THE BIOSYNTHESIS

OF ANGUILLABACTIN; AN IRON-BINDING SIDEROPHORE.

CC -1 SIMILARITY: TO OTHER THIOESTERASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL: M34504; AAA79861.1; -

DR EMBL: Z12000; CAA78045.1; -

DR PIR: J00417; J00417.

DR PIR: S26422; S26422.

DR InterPro: IPR01031; Thioesterase.

DR Pfam: PF00975; Thioesterase. 1.

KW Plasmid; Hydrolase.

FT ACT_SITE 92 BY SIMILARITY.

FT ACT_SITE 229 BY SIMILARITY.

SQ SEQUENCE 252 AA; 28070 MW; 1FB1AA3CCEDB99F4 CRC64;

Query Match 65.1%; Score 41; DB 1; Length 252;
Best Local Similarity 66.7%; Pred. No. 15;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SPRMWPCTC 9
11111
Db 156 SPRMWPCTC 164

RESULT 8
C132_MYCTU STANDARD: PRT: 461 AA.

ID C132_MYCTU STANDARD: PRT: 461 AA.

AC P77900;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Putative cytochrome P450 132 (EC 1.14.-.-)

GN CYP132 OR RV1394C OR MT1439 OR MTCY21B4.11C.

OS Mycobacterium tuberculosis.

OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;

Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1773;

RP SEQUENCE FROM N.A.

RX MEDLINE=96295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,

RA Horsby T., Jagsels K., Krogh A., McLean J., Mole S., Murphy L.,

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

RA Ruster S., Seeger K., Skelton S., Squares S., Squares R.,

RA Sutton J.E., Taylor K., Whitehead S., Barrett B.C.;

RT "Deciphering the biology of Mycobacterium tuberculosis from the

complete genome sequence.";

RL Nature 393:537-544(1998).

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=CDC 1551 / Oshkosh;

RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,

RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,

RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,

RA Delcher A., Utterback T., Weidman J., Knout H., Gill J., Mikula A.,

RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical
 RT laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBS databases.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: 280108; CAB02176.1; -
 DR EMBL: AE007015; AAK45704.1; -
 DR HSSP: P14779; IUPZ.
 DR TIGR: W11439; -
 DR Tuberculin; R1394C; -
 DR InterPro: IPR001128; Cytochrome_P450.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 DR Hypothetical protein; Oxidoreductase; Monooxygenase; Heme;
 KM Complete proteome.
 FT BINDING 409 HEME (BY SIMILARITY).
 FT CONFLICT 135 135 R -> L (IN REF. 2).
 FT SEQUENCE 461 AA; 52229 MW; 2DEF61C8A10E0CF3 CRC64;
 SQ
 Query Match 65.1%; Score 41; DB 1; Length 461;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 PRMWP 6
 Db 373 PRMWP 377
 ID ATP6_NAERO STANDARD; PRT; 119 AA.
 AC P22067;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ATP synthase A chain (EC 3.6.3.14) (Protein 6) (Fragment).
 GN ATP6 OR OL12.
 OS Naegleria fowleri.
 OG Mitochondrion.
 OC Eukaryota; Eukaryota; Schizopyrenida; Vahlkampfiidae; Naegleria.
 OX NCBI_TaxID=5763;
 RX MEDLINE=91178040; PubMed=2007628;
 RA McLaughlin G.L., Vodkin M.H., Huizinga H.W.;
 RT "Amplification of repetitive DNA for the specific detection of
 RT Naegleria fowleri.";
 RL J. Clin. Microbiol. 29:227-230(1991).
 CC -1- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL. IT MAY PLAY A
 CC DIRECT ROLE IN THE TRANSLLOCATION OF PROTONS ACROSS THE MEMBRANE.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
 CC H(+) (out).
 CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
 CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
 CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), Epsilon(1). CF(0)
 CC HAS THREE MAIN SUBUNITS: A, B AND C.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M55009; CAB25936.1; -
 DR PIR: A53257; A53257.
 DR InterPro: IPR000568; ATPsynth_Asub.
 DR Pfam: PF00119; ATP-synt_A; 1.
 DR PROSITE: PS00449; ATPASE_A; PARTIAL.
 KM Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
 FT NON_TER 1
 FT NON_TER 1
 FT SEQUENCE 119 AA; 13934 MW; 50892FC0B5C04F5 CRC64;
 SQ
 Query Match 61.9%; Score 39; DB 1; Length 119;
 Best Local Similarity 83.3%; Pred. No. 15;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 WMPYCL 9
 Db 70 WMPYCL 75
 ID ZRT1_YEAST STANDARD; PRT; 376 AA.
 AC P32804;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE ZRT1 protein.
 GN ZRT1 OR YGL255W OR NRC376.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RX MEDLINE=93311123; PubMed=8322518;
 RA Breitwieser W., Price C., Schuster T.;
 RT "Identification of a gene encoding a novel zinc finger protein in
 RT Saccharomyces cerevisiae.";
 RL Yeast 9:551-556(1993).
 GN Yeast 9:551-556(1993).
 GN [2]
 CC -1- FUNCTION: HIGH-AFFINITY ZINC TRANSPORT PROTEIN.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: TO YEAST YLR130C AND S. POMBE SPBC16D10.06.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X67787; CAA47997.1; -
 DR EMBL: X94357; CAA64132.1; -
 DR EMBL: Z72777; CAA96975.1; -
 DR PIR: S28553; S28553.
 DR PIR: S33654; S33654.
 DR SGD: S0003224; ZRT1.
 DR InterPro: IPR004698; ZIP_transport.
 DR InterPro: IPR003689; Zn_tnprt_Zip.

DR Pfam: PF02535; zip: 1.
 DR Transmembrane; TIGR00820; zip: 1.
 KM Transmembrane; Transport; Zinc.
 FT TRANSMEM 51 71 POTENTIAL.
 FT TRANSMEM 81 101 POTENTIAL.
 FT TRANSMEM 131 151 POTENTIAL.
 FT TRANSMEM 221 241 POTENTIAL.
 FT TRANSMEM 283 303 POTENTIAL.
 FT TRANSMEM 315 335 POTENTIAL.
 SQ SEQUENCE 376 AA; 41581 MW; 7A1F8367D49BAC3C CRC64;

Query Match
 Best Local Similarity 61.9%; Score 39; DB 1; Length 376;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SPRWMPCL 9
 DB 277 SKRWMPAL 285

RESULT 11
 YCCA_CAEEL STANDARD: PRT; 175 AA.
 ID YCCA_CAEEL
 AC Q23280;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 19.4 kDa protein ZC395.10 in chromosome III.
 GN ZC395.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Connell M.;
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: BELONGS TO THE P23 / W052 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U13642; AAG00038.1; -
 DR Wormpep; ZC395.10; CE01436.
 KM Hypothetical protein.
 FT DOMAIN 145 175 ASP/GLU-RICH.
 FT DOMAIN 165 168 POLY-GLU.
 SQ SEQUENCE 175 AA; 19431 MW; D5C136F30446E37A CRC64;

Query Match
 Best Local Similarity 60.3%; Score 38; DB 1; Length 175;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 SPRWMPCL 9
 DB 81 TPWWPRL 89

RESULT 12
 LYCH_CHASP STANDARD: PRT; 211 AA.
 ID LYCH_CHASP
 AC P00721;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE N,O-diacyltlumuramide (EC 3.2.1.-) (Lysozyme CH).
 OS Chalaropsis sp.

OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Chalaropsis.
 OX NCBI_TaxID=36534;
 RN (1)
 RP SEQUENCE.
 RX MEDLINE=75151523; PubMed=1168638;
 RA Felch J.W., Inagami T., Hash J.H.;
 RT "The N,O-diacyltlumuramide of Chalaropsis species. V. The complete
 RT amino acid sequence."
 RL J. Biol. Chem. 250:3713-3720(1975).
 RN [2]
 RP ACTIVE SITE.
 RX MEDLINE=79005662; PubMed=567645;
 RA Fouché P.B., Hash J.H.;
 RT "The N,O-diacyltlumuramide of Chalaropsis species. Identification of
 RT aspartyl and glutamyl residues in the active site."
 RL J. Biol. Chem. 253:6787-6793(1978).
 CC -1 FUNCTION: THIS EXTRACELLULAR ENZYME HAS BOTH LYSOZYME
 CC (ACETYLMURAMIDASE) AND DIACETYLMURAMIDASE ACTIVITIES.
 CC -1 CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages between N-
 CC acetyl-D-glucosamine and N-acetylmuramic acid in peptidoglycan
 CC heteropolymers of the prokaryotes cell walls.
 CC -1 SIMILARITY: BELONGS TO FAMILY 25 OF GLYCOSYL HYDROLASES.
 DR PIR: A00876; MKRAD.
 DR Interpro: IPR002053; GH_25.
 DR Pfam: PF01183; Glyco_Hydro_25; 1.
 DR Prodom: PD004620; GH_25; 1.
 DR PROSITE: PS00953; GLYCOSYL_HYDROL_F25; 1.
 KW Hydrolase; Glycosidase; Bacteriolytic enzyme.
 FT ACT_SITE 6 6
 FT ACT_SITE 33 33
 FT DISULFID 108 147
 SQ SEQUENCE 211 AA; 22413 MW; 379D758A383EC38C CRC64;

Query Match
 Best Local Similarity 60.3%; Score 38; DB 1; Length 211;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SPRWMPCL 8
 DB 140 NPSWSSC 147

RESULT 13
 OX11_SCHPO STANDARD: PRT; 374 AA.
 ID OX11_SCHPO
 AC O14300;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytochrome oxidase biogenesis protein oxal-1, mitochondrial precursor.
 GN OXAL-1 OR SPAC9G1.04.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=20177828; PubMed=10712694;
 RA Bonney N., Kermorgant M., Groudinsky O., Dujardin G.;
 RT "The respiratory gene OXAL has two fission yeast orthologues which
 RT together encode a function essential for cellular viability."
 RL Mol. Microbiol. 35:1135-1145(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouris J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabblowitsch E.,
 RA Rutherford K., Rutter S., Saunders S., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Filicz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Manbult R., Purnelle B.,
 RA Coffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.U., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.,
 RT The genome sequence of *Schizosaccharomyces pombe*.";
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: REQUIRED FOR A POST-TRANSLATIONAL STEP OF CYTOCHROME
 CC OXIDASE BIOGENESIS. SEEMS TO INVOLVED IN THE PROTEOLYTIC
 CC PROCESSING OF CYTOCHROME OXIDASE SUBUNIT 2. OXAL-2 IS ESSENTIAL
 CC FOR VIABILITY WHEREAS OXAL-1 IS NOT. WHEN BOTH ARE DELETED THE
 CC CELL IS NON-VIABLE, THEREFORE OXAL-1 ACT AS A BACK-UP FOR OXAL-2.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC (Probable).
 CC -1- SIMILARITY: BELONGS TO THE OXAL / 60 KDA IMP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X94123; CAA63843.1; -;
 DR EMBL: Z98763; CAB11488.1; -;
 DR InterPro: IPR001708; 60kDa_Innermem.
 DR Pfam: PF02096; 60kD_IMP; 1.
 KW Transmembrane; Mitochondrion; Multigene family; Transil peptide.
 FT TRANSIT 1 MITOCHONDRION (POTENTIAL).
 FT CHAIN ? 374 CYTOCHROME OXIDASE BIOGENESIS PROTEIN
 FT OXAL-1.
 FT TRANSMEM 77 97 POTENTIAL.
 FT TRANSMEM 160 180 POTENTIAL.
 FT TRANSMEM 202 222 POTENTIAL.
 FT TRANSMEM 243 263 POTENTIAL.
 SQ SEQUENCE 374 AA; 42060 MW; 989314857793402D CRC64;
 Query Match 60.3%; Score 38; DB 1; Length 374;
 Best Local Similarity 55.6%; Pred. No. 58;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SPRMPTCL 9
 Db 62 NPSMWPYAL 70
 RESULT 14
 METX_PSEAE STANDARD; PRT; 379 AA.
 ID METX_PSEAE
 AC P57714; O916A5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Homoserine O-acetyltransferase (EC 2.3.1.31) (Homoserine O-trans-
 DE acetylase) (Homoserine transacetylase) (HTA).
 GN METX OR PA0390.
 OS Pseudomonas aeruginosa.
 CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltz L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.K., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Mong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lofy S., Olson M.V., an
 RT Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + L-homoserine -> CoA + O-acetyl-L-
 CC homoserine.
 CC -1- PATHWAY: Methionine biosynthesis; HTA variant; first step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE AB HYDROLASE FAMILY. HTA SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE004476; MAG03779.1; -;
 DR InterPro: IPR000073; Abhydrolase.
 DR Pfam: PF00561; abhydrolase; 1.
 KW Methionine biosynthesis; Transferase; Acyltransferase;
 KW Complete proteome.
 FT ACT_SITE 157 157 POTENTIAL.
 FT ACT_SITE 356 356 POTENTIAL.
 SQ SEQUENCE 379 AA; 41834 MW; 8744551ABF35AF45 CRC64;
 Query Match 60.3%; Score 38; DB 1; Length 379;
 Best Local Similarity 50.0%; Pred. No. 59;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 PRMPTCL 9
 Db 76 PGWMDSCI 83
 RESULT 15
 O85E_DROME STANDARD; PRT; 417 AA.
 AC P81924; O9VHP5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Odorant receptor 85e.
 GN O85E OR OR104 OR DOR104 OR CG9700.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
 RC STRAIN=Oregon-R; Tissue=Maxillary palps;
 RX MEDLINE=99189757; PubMed=10089887;
 RA Vossell L.B., Amrein H., Morozov P.S., Rhetsky A., Axel R.,
 RT "A spatial map of olfactory receptor expression in the *Drosophila*
 RL cell 96:725-736(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner K.A., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Abayanti A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale U., Bayraktiroglu U., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borrova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu L.B., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Decher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dopp L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.D., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jamal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pitman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
CC -I- FUNCTION: PROBABLE ROLE IN THE ODORANT RESPONSE, BEING AN ODORANT
CC RECEPTOR.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -I- TISSUE SPECIFICITY: EXPRESSED IN 15% OF THE 120 SENSORY NEURONS
CC WITHIN THE MAXILLARY PALP.
CC -I- SIMILARITY: BELONGS TO FAMILY DR-OR OF G-PROTEIN COUPLED
CC RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF127922; AAD26357.1; -;
DR EMBL: AE003679; AAF54256.1; -;
DR FlyBase: FBgn0026399; O885e.
DR InterPro: IPR004117; 7tm.6.
DR Pfam: PF02949; 7tm.6; 1.
KW Transmembrane; G-protein coupled receptor; Olfaction; Glycoprotein;
KW Multigene family.
FT DOMAIN 1 60 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 61 81 1 (POTENTIAL).
FT DOMAIN 82 98 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 99 119 2 (POTENTIAL).
FT DOMAIN 120 159 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 160 180 3 (POTENTIAL).
FT DOMAIN 181 212 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 213 233 4 (POTENTIAL).
FT DOMAIN 234 286 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 287 307 5 (POTENTIAL).
FT DOMAIN 308 334 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 335 355 6 (POTENTIAL).
FT DOMAIN 356 367 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 368 388 7 (POTENTIAL).
FT DOMAIN 389 417 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 39 39 N-LINKED (GLCNC. . .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNC. . .) (POTENTIAL).
FT CONFLICT 410 417 NVCSTPK -> FTRDIIIFLYNSRAVHTAGFRYYMDV
FT NRKSVITQAFSLFTLLQAKKTESEL (IN REF.
FT 1).
SQ SEQUENCE 417 AA; 47212 MW; A27C0A9A76040C70 CRC64;
Query Match 60.3%; Score 38; DB 1; Length 417;
Best Local Similarity 80.0%; Pred. No. 64;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 2 PRMMP 6
DB 48 PKMWP 52

Search completed: March 27, 2003, 16:11:10
Job time : 3.52632 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 16:01:02 ; Search time 2.63158 Seconds
(without alignments)
704.681 Million cell updates/sec

Title: US-10-019-219-2

Perfect score: 63

Sequence: 1 SPRMWPRL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	46	73.0	806	10	09ZW71	09ZW71 arabidopsis
2	45	71.4	136	10	09XF33	09XF33 oryza sativ
3	43	68.3	136	10	09X0S8	09X0S8 rosa hybrid
4	43	68.3	139	2	09EX86	09EX86 planobispor
5	43	68.3	145	12	09Z977	09Z977 hepatitis c
6	43	68.3	145	12	09W895	09W895 hepatitis c
7	43	68.3	145	12	072123	072123 hepatitis c
8	43	68.3	145	12	072124	072124 hepatitis c
9	43	68.3	145	12	072125	072125 hepatitis c
10	43	68.3	145	12	072126	072126 hepatitis c
11	43	68.3	237	4	09H7W0	09H7W0 homo sapien
12	43	68.3	383	10	09FMG2	09FMG2 oryza sativ
13	43	68.3	458	10	09FRJ1	09FRJ1 oryza sativ
14	43	68.3	615	10	0948P3	0948P3 cucumis mel
15	43	68.3	3010	12	09Z969	09Z969 hepatitis c
16	43	68.3	3010	12	09Z970	09Z970 hepatitis c

17	43	68.3	3010	12	09Z971	09Z971 hepatitis c
18	43	68.3	3010	12	09Z972	09Z972 hepatitis c
19	43	68.3	3010	12	00Z828	00Z828 h genome po
20	42	66.7	136	5	025300	025300 leishmania
21	42	66.7	221	16	09JZ46	09JZ46 neisseria m
22	42	66.7	221	16	09JU40	09JU40 neisseria m
23	42	66.7	377	10	09XEP9	09XEP9 sorghum bic
24	42	66.7	577	5	096756	096756 dugesia tlg
25	42	66.7	858	5	027681	027681 leishmania
26	41	65.1	103	12	09Q2P0	09Q2P0 human herpe
27	41	65.1	123	17	09YDA5	09YDA5 aeropyrum p
28	41	65.1	152	16	09L274	09L274 streptomyce
29	41	65.1	207	5	08STY1	08STY1 encephalito
30	41	65.1	224	16	093503	093503 streptomyce
31	41	65.1	295	16	08YCI7	08YCI7 bruceella me
32	41	65.1	629	10	08S1A6	08S1A6 oryza sativ
33	41	65.1	790	5	020599	020599 caenorhabdl
34	41	65.1	842	5	0950P5	0950P5 caenorhabdl
35	40	63.5	64	2	007030	007030 vibrio chol
36	40	63.5	169	11	09D0G4	09D0G4 mus musculu
37	40	63.5	169	11	09CX23	09CX23 mus musculu
38	40	63.5	169	11	09RLR6	09RLR6 mus musculu
39	40	63.5	212	16	055157	055157 synecocyst
40	40	63.5	253	2	051552	051552 pseudomonas
41	40	63.5	274	16	098EP8	098EP8 rhizobium l
42	40	63.5	290	2	003075	003075 bradyrhizob
43	40	63.5	319	2	0931F9	0931F9 propionibac
44	40	63.5	403	2	093PL2	093PL2 treponema d
45	40	63.5	471	10	09LX16	09LX16 arabidopsis

ALIGNMENTS

RESULT 1

ID	Query Match	Length	DB	ID	Description
09ZW71	73.0	806	10	09ZW71	09ZW71 arabidopsis
AC	71.4	136	10	09XF33	09XF33 oryza sativ
DT	68.3	136	10	09X0S8	09X0S8 rosa hybrid
DE	68.3	139	2	09EX86	09EX86 planobispor
GN	68.3	145	12	09Z977	09Z977 hepatitis c
OS	68.3	145	12	09W895	09W895 hepatitis c
OC	68.3	145	12	072123	072123 hepatitis c
OC	68.3	145	12	072124	072124 hepatitis c
OC	68.3	145	12	072125	072125 hepatitis c
OC	68.3	145	12	072126	072126 hepatitis c
OX	68.3	237	4	09H7W0	09H7W0 homo sapien
RN	68.3	383	10	09FMG2	09FMG2 oryza sativ
RP	68.3	458	10	09FRJ1	09FRJ1 oryza sativ
RC	68.3	615	10	0948P3	0948P3 cucumis mel
RX	68.3	3010	12	09Z969	09Z969 hepatitis c
RA	68.3	3010	12	09Z970	09Z970 hepatitis c
RA	68.3	3010	12	09Z971	09Z971 hepatitis c
RA	68.3	3010	12	09Z972	09Z972 hepatitis c
RA	68.3	3010	12	00Z828	00Z828 h genome po
RA	68.3	3010	12	025300	025300 leishmania
RA	68.3	3010	12	09JZ46	09JZ46 neisseria m
RA	68.3	3010	12	09JU40	09JU40 neisseria m
RA	68.3	3010	12	09XEP9	09XEP9 sorghum bic
RA	68.3	3010	12	096756	096756 dugesia tlg
RA	68.3	3010	12	027681	027681 leishmania
RA	68.3	3010	12	09Q2P0	09Q2P0 human herpe
RA	68.3	3010	12	09YDA5	09YDA5 aeropyrum p
RA	68.3	3010	12	09L274	09L274 streptomyce
RA	68.3	3010	12	08STY1	08STY1 encephalito
RA	68.3	3010	12	093503	093503 streptomyce
RA	68.3	3010	12	08YCI7	08YCI7 bruceella me
RA	68.3	3010	12	08S1A6	08S1A6 oryza sativ
RA	68.3	3010	12	020599	020599 caenorhabdl
RA	68.3	3010	12	0950P5	0950P5 caenorhabdl
RA	68.3	3010	12	007030	007030 vibrio chol
RA	68.3	3010	12	09D0G4	09D0G4 mus musculu
RA	68.3	3010	12	09CX23	09CX23 mus musculu
RA	68.3	3010	12	09RLR6	09RLR6 mus musculu
RA	68.3	3010	12	055157	055157 synecocyst
RA	68.3	3010	12	051552	051552 pseudomonas
RA	68.3	3010	12	098EP8	098EP8 rhizobium l
RA	68.3	3010	12	003075	003075 bradyrhizob
RA	68.3	3010	12	0931F9	0931F9 propionibac
RA	68.3	3010	12	093PL2	093PL2 treponema d
RA	68.3	3010	12	09LX16	09LX16 arabidopsis

Query Match: 73.0% Score 46; DB 10; Length 806;
Best Local Similarity: 75.0%; Pred. No. 30;
Matches: 6; Conservative: 1; Mismatches: 1; Indels: 0; Gaps: 0;

OY 2 PRMWPTCL 9
1:|||||
DB 676 PRMWPTSL 683

RESULT 2

OY 09XF33 PRELIMINARY: PRT: 717 AA.
AC 09XF33;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
DE Hypothetical 79.0 kDa protein.
OS Oryza sativa (indica cultivar group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39946;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. TEQING;
RA Liaca V., Lou A., Young S., Messing J.;
RT "Microcollinearity in cereal genomes."
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF128457; AAD27632.1; -;
DR InterPro: IPR00210; BTB_POZ.
DR Pfam: PF00651; BTB; 1.
DR SMART: SM00225; BTB; 1.
DR PROSITE: PS50097; BTB; 1.
KW Hypothetical protein
SQ SEQUENCE 717 AA; 79014 MW; 90DB888B78119E60 CRC64;

Query Match 71.4%; Score 45; DB 10; Length 717;
Best Local Similarity 71.4%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 PRMWPTC 8
1:|||||
DB 302 PRMWPTC 308

RESULT 3

ID 0940S8 PRELIMINARY: PRT: 136 AA.
AC 0940S8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE EIN3-like transcription factor (Fragment).
GN EIN3.
OS Rosa hybrid cultivar.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Euposids I; Rosales; Rosaceae; Rosoideae; Rosa.
OX NCBI_TaxID=128735;
RN [1]
RP SEQUENCE FROM N.A.
RA Mueller R., Owen C.A., Stummann B.M.;
RT "Partial sequence of Rosa hybrida cultivar mRNA for EIN3-like
transcription factor."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY052825; AAL14267.1; -;
FT NON_TER 1
FT NON_TER 136
SQ SEQUENCE 136 AA; 15267 MW; B77FAD5FB81383B9 CRC64;

Query Match 68.3%; Score 43; DB 10; Length 136;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SPRMWPT 7
1:|||||

DB 106 SPRMWPT 112

RESULT 4

ID 09EX86 PRELIMINARY: PRT: 139 AA.
AC 09EX86;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Putative peptide synthetase (Fragment).
OS Planobispora rosea.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptosporangineae; Streptosporangiaceae;
OC Planobispora.
OX NCBI_TaxID=35762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 53733;
RX MEDLINE=20535709; PubMed=11085259;
RA Sosio M., Bossi E., Bianchi A., Donadio S.;
RT "Multiple peptide synthetase gene clusters in actinomycetes."
RL Mol. Gen. Genet. 264:213-221(2000).
DR EMBL: AF276363; CAC01622.1; -;
DR HSP; P14687; IAMU.
DR InterPro: IPR000873; AMP_bind.
DR InterPro: IPR003880; Pantine_attach.
DR Pfam: PF00501; AMP-binding; 1.
DR Pfam: PF00550; pp-binding; 1.
DR PROSITE: PS50075; ACP_DOMAIN; 1.
FT NON_TER 139
FT NON_TER 139
SQ SEQUENCE 139 AA; 15029 MW; 1F2489785FD715C6 CRC64;

Query Match 68.3%; Score 43; DB 2; Length 139;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SPRMWPT 7
1:|||||
DB 31 SPRMWPT 37

RESULT 5

ID 092977 PRELIMINARY: PRT: 145 AA.
AC 092977; 092978;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788;
RA Yanagi M., St. Claire M., Shapiro M., Emerson S.U., Purcell R.H.,
Bukh J.;
RT "Transcripts of a chimeric cDNA clone of hepatitis C virus genotype 1b
are infectious in vivo."
RL Virology 244:161-172(1998).
DR EMBL: AF054260; AAC15733.1; -;
DR EMBL: AF054261; AAC15734.1; -;
DR EMBL: AF054265; AAC15738.1; -;
DR InterPro: IPR002166; HCV_RdRP.
DR Pfam: PF00998; HCV_RdRP; 1.
KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
FT NON_TER 1
FT NON_TER 145
SQ SEQUENCE 145 AA; 16235 MW; 8157D290205C2252 CRC64;

Query Match 68.3%; Score 43; DB 12; Length 145;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRMWPCTCL 9
DB 123 PRMWPCTCL 130

RESULT 6

09W895 PRELIMINARY; PRT; 145 AA.
AC 09W895;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Genome polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric cDNA clone of hepatitis C virus genotype 1b,
RT are infectious in vivo."
RL Virology 244:161-172(1998).
DR EMBL; AF054267; AAC15740.1; -;
DR EMBL; AF054263; AAC15736.1; -;
DR InterPro; IPR002166; HCV_RdRp.
DR Pfam; PF00998; HCV_RdRp; 1.
KW Nonstructural protein; Polypeptide; RNA-directed RNA polymerase.
FT NON_TER 1 145
SQ SEQUENCE 145 AA; 16175 MW; 8157D280215C2252 CRC64;

Query Match 68.3%; Score 43; DB 12; Length 145;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRMWPCTCL 9
DB 123 PRMWPCTCL 130

RESULT 7

072123 PRELIMINARY; PRT; 145 AA.
AC 072123;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Genome polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
RT infectious in vivo."
RL Virology 244:161-172(1998).
DR EMBL; AF054262; AAC15735.1; -;
DR InterPro; IPR002166; HCV_RdRp.
DR Pfam; PF00998; HCV_RdRp; 1.

KW Nonstructural protein; Polypeptide; RNA-directed RNA polymerase.
FT NON_TER 1 145
SQ SEQUENCE 145 AA; 16145 MW; 8157D29C3C9DE252 CRC64;

Query Match 68.3%; Score 43; DB 12; Length 145;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRMWPCTCL 9
DB 123 PRMWPCTCL 130

RESULT 8

072124 PRELIMINARY; PRT; 145 AA.
AC 072124;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Genome polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
RT infectious in vivo."
RL Virology 244:161-172(1998).
DR EMBL; AF054264; AAC15737.1; -;
DR InterPro; IPR002166; HCV_RdRp.
DR Pfam; PF00998; HCV_RdRp; 1.
KW Nonstructural protein; Polypeptide; RNA-directed RNA polymerase.
FT NON_TER 1 145
SQ SEQUENCE 145 AA; 16087 MW; 83A9FD4C3C9DE252 CRC64;

Query Match 68.3%; Score 43; DB 12; Length 145;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRMWPCTCL 9
DB 123 PRMWPCTCL 130

RESULT 9

072125 PRELIMINARY; PRT; 145 AA.
AC 072125;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Genome polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
RT infectious in vivo."
RL Virology 244:161-172(1998).

DR EMBL: AF054266; AAC15739.1; -
DR InterPro: IPR002166; HCV_RdRp.
DR Pfam: PF00998; HCV_RdRp. 1.
KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
FT NON_TER 1
FT NON_TER 145
SQ SEQUENCE 145 AA; 16175 MW; 8157C2CD7999E252 CRC64;

Query Match 68.3%; Score 43; DB 12; Length 145;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 PRMPTCL 9
DB 123 PRMPTCL 130

RESULT 10

ID 072126 PRELIMINARY; PRT; 145 AA.
AC 072126;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HC-24;
RX MEDLINE=98240944; PubMed=9581788;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
RT infectious in vivo.";
RL Virology 244:161-172(1998).
DR EMBL: AF054266; AAC15741.1; -
DR InterPro: IPR002166; HCV_RdRp.
DR Pfam: PF00998; HCV_RdRp. 1.
KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
FT NON_TER 1
FT NON_TER 145
SQ SEQUENCE 145 AA; 16175 MW; 8157C2CD7999E252 CRC64;

Query Match 68.3%; Score 43; DB 12; Length 145;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 PRMPTCL 9
DB 123 PRMPTCL 130

RESULT 11

ID 09H7W0 PRELIMINARY; PRT; 237 AA.
AC 09H7W0;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CDNA FLJ14202 f1s, clone NT2RP3002985.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagaatsuma M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,

RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
RA Niimiya K., Iwayanagi T.;
RT "NEO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AK024264; BAB14864.1; -
SQ SEQUENCE 237 AA; 25615 MW; 2CE39E9D320B3863 CRC64;

Query Match 68.3%; Score 43; DB 4; Length 237;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 WMPPTC 8
DB 99 WMPPTC 103

RESULT 12

ID 09FWG2 PRELIMINARY; PRT; 383 AA.
AC 09FWG2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical 41.9 kDa protein.
GN OSJNB001511.26.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=CV. NIPONBARE;
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
RA Hsiao J., Zismann V., Pal G., Bowman C.L., Fujii C.Y., Vanaken S.E.,
RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSJNB001511 genomic sequence.";
RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AC051633; AAG13599.1; -
DR HSP: P35813; 1A60.
DR InterPro: IPR000222; PP2C.
DR InterPro: IPR001932; PP2C-like.
DR Pfam: PF00481; PP2C. 1.
DR SMART: SM00332; PP2Cc. 1.
DR SMART: SM00331; PP2C_SIG. 1.
DR PROSITE: PS01032; PP2C. 1.
KW Hypothetical protein.

SQ SEQUENCE 383 AA; 41899 MW; 9B58D7BCD5217B81 CRC64;

Query Match 68.3%; Score 43; DB 10; Length 383;
Best Local Similarity 71.4%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 RWMPTCL 9
DB 14 RWMPTCL 20

RESULT 13

ID 09FRJ1 PRELIMINARY; PRT; 458 AA.
AC 09FRJ1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical 49.5 kDa protein.
GN OSJNB0064P21.12.
OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.

OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Burr P.C., Hsiao J.,
 Zismann V., Pal G., Bowman C.L., Fujii C.Y., Vanhaken S.E.,
 Bowman C.L., Craven B., Utterback T.R., Kraik H., Feldblum T.V.,
 Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 10 BAC OSJNB0064P21 genomic sequence.";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC073166; AAC46118.1; -
 DR HSSP: P35813; 1A60.
 DR InterPro: IPR000222; PP2C.
 DR InterPro: IPR001932; PP2C-like.
 DR Pfam: PF00481; PP2C; 1.
 DR SMART: SM00332; PP2C; 1.
 DR SMART: SM00331; PP2C-SIG; 1.
 DR PROSITE: PS01032; PP2C; 1.
 KM Hypothetical protein
 SQ SEQUENCE 458 AA; 49485 MW; C7A857C561DA2B50 CRC64;

Query Match 68.3%; Score 43; DB 10; Length 458;
 Best Local Similarity 71.4%; Pred. No. 52;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PRMPTCL 9
 DB 14 RMPPTCV 20

RESULT 14
 Q948P3 PRELIMINARY; PRT; 615 AA.
 AC Q948P3;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE EIN3-like protein.
 GN CMEI12.
 OS Cucumis melo (Muskmelon).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids I; Cucurbitales; Cucurbitaceae; Cucumis.
 OX NCBI_TaxID=3636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. ANDES;
 RA Takekawa K., Sawaki T., Parlasca J., Matsumura A., Fujimori A.,
 Tatsuno Y., Asama H., Sonoda M., Hirabayashi T., Nakagawa H., Sato T.;
 RT "Cloning of cDNA encoding melon EIN3-like genes (CmeI1 and 2) from
 ripening fruit."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB063192; BAB64345.1; -
 KM DNA-binding; Zinc-finger
 SQ SEQUENCE 615 AA; 70183 MW; 11BB3C5A7246EFP0 CRC64;

Query Match 68.3%; Score 43; DB 10; Length 615;
 Best Local Similarity 85.7%; Pred. No. 66;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SPRMPT 7
 DB 212 SPMWPT 218

RESULT 15
 O92969 PRELIMINARY; PRT; 3010 AA.
 AC O92969;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Genome polypeptide [Contains: envelope glycoprotein E2 (GP68) (GP70)

DE (NS1).
 OC Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HC-J4;
 RX MEDLINE=98240944; PubMed=9581788;
 RA Yanagi M., St. Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
 Bukh J.;
 RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
 infectious in vivo."
 RL Virology 244:161-172(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HC-J4;
 RA Yanagi M., Bukh J.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
 DR EMBL: AF054247; AAC15722.1; -
 DR HSSP: P26663; 1JXP.
 DR MEROPS: S29.001; -
 DR DR MEROPS: U39.001; -
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR002522; HCV_capsid.
 DR InterPro: IPR002521; HCV_core.
 DR InterPro: IPR002519; HCV_env.
 DR InterPro: IPR002531; HCV_NS1.
 DR InterPro: IPR002518; HCV_NS2.
 DR InterPro: IPR004109; HCV_NS3.
 DR InterPro: IPR000745; HCV_NS4a.
 DR InterPro: IPR001490; HCV_NS4b.
 DR InterPro: IPR002868; HCV_NS5a.
 DR InterPro: IPR002166; HCV_RdRp.
 DR Pfam: PF01543; HCV_capsid; 1.
 DR Pfam: PF01542; HCV_core; 1.
 DR Pfam: PF01539; HCV_env; 1.
 DR Pfam: PF01560; HCV_NS1; 1.
 DR Pfam: PF01538; HCV_NS2; 1.
 DR Pfam: PF02907; HCV_NS3; 1.
 DR Pfam: PF01006; HCV_NS4a; 1.
 DR Pfam: PF01001; HCV_NS4b; 1.
 DR Pfam: PF01506; HCV_NS5a; 1.
 DR Pfam: PF00998; HCV_RdRp; 1.
 DR Pfam: PF0186062; HCV_NS1; 1.
 KW Amp-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
 KW Nonstructural protein; Polypeptide; RNA-directed RNA polymerase;
 KW Transmembrane.
 SQ SEQUENCE 3010 AA; 326776 MW; 9B3FD910CF00E2C5 CRC64;

Query Match 68.3%; Score 43; DB 12; Length 3010;
 Best Local Similarity 75.0%; Pred. No. 2.4e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRMPTCL 9
 DB 2988 PRMPTCL 2995

Search completed: March 27, 2003, 16:12:06
 Job time: 3.63158 secs

THIS PAGE BLANK (USPTO)